# Multiplicativity of completely bounded p-norms implies a new additivity result

Igor Devetak
Electrical Engineering Department, University of Southern California, Los Angeles CA 90089
devetak@csi.usc.edu

Marius Junge
Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana IL 61801
mjunge@math.uiuc.edu

Christopher King
Department of Mathematics, Northeastern University, Boston MA 02115
king@neu.edu

Mary Beth Ruskai
Department of Mathematics, Tufts University, Medford MA 02155
Marybeth.Ruskai@tufts.edu

February 9, 2008

**Abstract**

We prove additivity of the minimal conditional entropy associated with a quantum channel $\Phi$, represent by a completely positive (CP), trace-preserving map, when the infimum of $S(\gamma_{12}) - S(\gamma_1)$ is restricted to states of the form $(\mathcal{I} \otimes \Phi)\big(|\psi\rangle\langle\psi|\big)$ We show that this follows from multiplicativity of the completely bounded norm of $\Phi$ considered as a map from $L_1 \to L_p$ for $L_p$ spaces defined by the Schatten p-norm on matrices, and give another proof based on entropy inequalities. Several related multiplicativity results are discussed and proved. In particular, we show that both the usual $L_1 \to L_p$ norm of a CP map and the corresponding completely bounded norm are achieved for positive semi-definite matrices. Physical interpretations are considered, and a new proof of strong subadditivity is presented.

# Contents

## 1 Introduction

Quantum channels are represented by completely positive, trace preserving (CPT) maps on $M_d$, the space of $d \times d$ matrices. Results and conjectures about additivity and superadditivity of various types of capacity play an important role in quantum information theory.

In this paper, we present a new additivity result which can be stated in terms of a type of minimal conditional entropy defined as

$$S_{\mathrm{CB,min}}(\Phi) = \inf_{\psi \in \mathbf{C}^d \otimes \mathbf{C}^d} \Big( S\big[(\mathcal{I} \otimes \Phi)\big(|\psi\rangle\langle\psi|\big)\big] - S\big[(\mathrm{Tr}_2\big(|\psi\rangle\langle\psi|\big)\big]\Big) \tag{1.1}$$

where $S(Q) = -\mathrm{Tr}\, Q \log Q$ is the von Neumann entropy. The shorthand CB stands for "completely bounded" which will be explained later. We will show that this CB minimal conditional entropy is additive, i.e.,

$$S_{\mathrm{CB,min}}(\Phi_A \otimes \Phi_B) = S_{\mathrm{CB,min}}(\Phi_A) + S_{\mathrm{CB,min}}(\Phi_B). \tag{1.2}$$

The expression (1.1) for $S_{\mathrm{CB,min}}(\Phi)$ should be compared to those for two important types of capacity. To facilitate this, it is useful to let $\gamma_{12} = (\mathcal{I} \otimes \Phi)\big(|\psi\rangle\langle\psi|\big)$, and observe that its reduced density matrices are $\gamma_1 = \mathrm{Tr}_2(\mathcal{I} \otimes \Phi)\big(|\psi\rangle\langle\psi|\big)$, and $\gamma_2 = \mathrm{Tr}_1(\mathcal{I} \otimes \Phi)\big(|\psi\rangle\langle\psi|\big)$. We can now rewrite (1.1) as

$$-S_{\mathrm{CB,min}}(\Phi) \;=\; \sup_{\psi} \Big[S(\gamma_1) - S(\gamma_{12})\Big]. \tag{1.3}$$

The capacity of a quantum channel for transmission of classical information when assisted by unlimited entanglement (as in, e.g., dense coding) is given by [5, 6, 16]

$$C_{EA}(\Phi) \;=\; \sup_{\psi} \Big[S(\gamma_1) + S(\gamma_2) - S(\gamma_{12})\Big]. \tag{1.4}$$

The capacity for transmission of quantum information without additional resources is the coherent information, [4, 10, 34, 46]

$$C_Q(\Phi) \;=\; \sup_{\psi} \Big[S(\gamma_2) - S(\gamma_{12})\Big] \tag{1.5}$$

In these expressions, the supremum is taken over all normalized vectors $\psi$ in $\mathbf{C}^d \otimes \mathbf{C}^d$ and $\gamma_{12}$ depends on both $\psi$ and $\Phi$. It has been established that $C_{EA}(\Phi)$ is additive [5, 16],

but $C_Q(\Phi)$ is not additive in general [11]. To understand the difference between $C_Q(\Phi)$ and $S_{\rm CB,min}(\Phi)$, use the trace-preserving property of $\Phi$ to rewrite $\gamma_1 = \text{Tr}_2\big(|\psi\rangle\langle\psi|\big)$ and $\gamma_2 = \Phi\big[\text{Tr}_1\big(|\psi\rangle\langle\psi|\big)\big]$. The additive quantity (1.3) contains $\gamma_1$ which is independent of $\Phi$, while the non-additive quantity (1.5) contains $\gamma_2$ which depends upon $\Phi$.

We do not have a completely satisfactory physical interpretation of the the CB entropy, although an operational meaning can be found. It appears to provide a measure of how well a channel preserves entanglement. In particular, if $\Phi$ is entanglement breaking, $S_{\rm CB,min}(\Phi) > 0$ (although the converse does not hold). Recently, Horodecki, Oppenheim and Winter [18] gave an elegant interpretation of quantum conditional information which we discuss in the context of our results in Section 5.

The additivity (1.2) will follow from the multiplicativity (2.5) of the quantity

$$\omega_p(\Phi) \equiv \sup_{\psi \in \mathbf{C}^d \otimes \mathbf{C}^d} \frac{\|(\mathcal{I} \otimes \Phi)\big(|\psi\rangle\langle\psi|\big)\|_p}{\|\text{Tr}_2\big(|\psi\rangle\langle\psi|\big)|_p} = \sup_{\psi \in \mathbf{C}^d \otimes \mathbf{C}^d} \frac{\|\gamma_{12}\|_p}{\|\gamma_1\|_p} \,. \tag{1.6}$$

We will see that this is a type of CB norm. Recall that one of several equivalent criteria for a map $\Phi$ to be completely positive is that for all integers $d$, the map $\mathcal{I}_d \otimes \Phi$ takes positive semi-definite matrices to positive semi-definite matrices. (We use $\mathcal{I}$ to denote the identity map $\mathcal{I}(\rho) = \rho$ to avoid confusion with the identity matrix I.) One can similarly define other concepts, such as completely isometric, in terms of the maps $\mathcal{I}_d \otimes \Phi$. The completely bounded (CB) norm is thus

$$\|\Phi\|_{\rm CB} = \sup_d \|\mathcal{I}_d \otimes \Phi\|. \tag{1.7}$$

However, this depends on the precise definition of the norm on the right side of (1.7) or, equivalently, on the norms used to regard $\Phi$ and $\mathcal{I}_d \otimes \Phi$ as maps between Banach spaces. The appropriate definitions for the situations considered here are described in Sections 3.1 and 3.3.

In the process of deriving our results, we obtain a number of related results of independent interest. For example, we show that when $\Phi$ is a CP map, both $\|\Phi\|_{q\to p}$ and the corresponding CB norm are attained for a positive semi-definite matrix, extending a result in [51]. The strong subadditivity (SSA) inequality [32, 42] for quantum entropy

$$S(Q_{123}) + S(Q_3) \leq S(Q_{23}) + S(Q_{13}) \tag{1.8}$$

is the basis for Holevo's proof of additivity of $C_{EA}(\Phi)$ and the proof of (1.2) given in Section 2.3. In Section 6 we use operator space methods to obtain a new proof of SSA.

This paper is organized as follows. Section 2 is concerned with our main result, (1.2). After some background, we present two different proofs. In Section 3, which is divided into six subsections, we introduce notation and summarize results about CB norms and operator spaces used in the paper. Only the basic notation in Section 3.1 and the Minkowski inequalities in Section 3.4 are needed for the main result, Theorem 11. A subtle distinction between the norms used to define $\|\Phi\|_{\rm CB}$ and $\|\mathcal{I}_d \otimes \Phi\|_{q\to p}$ often used in quantum information (e.g., [3, 26, 27, 51]) is described in the penultimate paragraph of Section 3.2.

In Section 4, we prove multiplicativity of the CB norm for maps $\Phi : L_q(M_m) \mapsto L_p(M_n)$. When $q \geq p$, we also show that the CB norm equals $\|\Phi\|_{q\to p}$, yielding a proof of multiplicativity for the latter. In Section 5, we explicitly give $\|\Phi\|_{\rm CB}$ and $S_{\rm CB,min}(\Phi)$ for simple examples, including the depolarizing channel; prove that $S_{\rm CB,min}(\Phi) > 0$ for EBT maps; and discuss physical interpretations. In Section 6, we use the Minkowski inequalities for the CB norms to obtain a new proof of SSA. We also show that the minimizer implicit in $\|X_{12}\|_{(1,p)}$ converges to $X_1$.

# 2 Additivity of CB entropy

## 2.1 Multiplicativity questions in quantum information theory

We are interested in CB norms when $\Phi$ is a map $L_q(M_d) \mapsto L_p(M_d)$ where $L_p(M_d)$ denotes the Banach space of $d \times d$ matrices with the Schatten norm $\|A\|_p = \big(\mathrm{Tr}\,|A|^p\big)^{1/p}$. One then defines the norm

$$\|\Phi\|_{q\to p} \equiv \sup_A \frac{\|\Phi(A)\|_p}{\|A\|_q} \tag{2.1}$$

Watrous [51] and Audenaert [2] independently showed that this norm is unchanged if the supremum in (2.1) is restricted to positive semi-definite matrices, resolving a question raised in [26]. Thus,

$$\|\Phi\|_{q\to p} = \sup_{A>0} \frac{\|\Phi(A)\|_p}{\|A\|_q} \tag{2.2}$$

In quantum information theory, the norm $\nu_p(\Phi) = \|\Phi\|_{1\to p}$ plays an important role. It has been conjectured [3] (see also [26]) that

$$\nu_p(\Phi_A \otimes \Phi_B) = \nu_p(\Phi_A)\, \nu_p(\Phi_B) \tag{2.3}$$

in the range $1 \leq p \leq 2$. Proof of this conjecture would imply additivity of minimal entropy which has been shown to be equivalent to several other important and long-standing conjectures [47]. We note here only that $S_{\min}(\Phi) = \inf_{\rho \in \mathcal{D}} S[\Phi(\rho)]$ where $\mathcal{D} = \{\rho : \rho > 0,\ \mathrm{Tr}\, \rho = 1\}$ denotes the set of density matrices. Note that $\nu_p(\Phi) = \sup_{\rho \in \mathcal{D}} \|\Phi(\rho)\|_p$. Amosov, Holevo and Werner [3] showed that the additivity of minimal entropy

$$S_{\min}(\Phi_A \otimes \Phi_B) = S_{\min}(\Phi_A) + S_{\min}(\Phi_B) \tag{2.4}$$

would follow if (2.3) can be proved.

In this paper, we consider instead $\|\Phi\|_{CB,1\to p}$ for which the expression in (1.7) reduces to $\omega_p(\Phi)$, and show that it is multiplicative, i.e., that

$$\omega_p(\Phi_A \otimes \Phi_B) = \omega_p(\Phi_A)\, \omega_p(\Phi_B). \tag{2.5}$$

We first show that (2.5) implies our new additivity result, providing a motivation for the technical material needed to prove (2.5). We subsequently found another proof which does not use CB norms; this is presented in Section 2.3. However, the CB proof given next provides an indication of the potential of this machinery for quantum information.

## 2.2 Proof of additivity from CB multiplicativity

We define a function of a self adjoint matrix with spectral decomposition $A = \sum_k \lambda_k |\phi_k\rangle\langle\phi_k|$ as $f(A) = \sum_k f(\lambda_k)|\phi_k\rangle\langle\phi_k|$. We will need functions of the form $f(t) = t^p \log t$ defined on $[0, \infty)$ so that $f(0) = 0$ for $p > 0$ and $Q^p \log Q$ is 0 on $\ker(Q)$. For any $Q > 0$ we define the entropy as $S(Q) = -\mathrm{Tr}\, Q \log Q$ and note that $S\big(\frac{Q}{\mathrm{Tr}\, Q}\big) = \frac{1}{\mathrm{Tr}\, Q} S(Q) + \log \mathrm{Tr}\, Q$.

We will often use the notation $\gamma_{12}$ for density matrices in the tensor product $M_d \otimes M_n \simeq M_{dn}$ and $\gamma_1 = \mathrm{Tr}_2\, \gamma_{12}$, for the corresponding reduced density matrix in $M_d$. (The partial trace $\mathrm{Tr}_2$ denotes the trace on $M_n$. One can similarly define $\gamma_2 = \mathrm{Tr}_1\, \gamma_{12}$ The density matrix $\gamma_{12}$ can be regarded as a probability distribution on $\mathbf{C}_d \otimes \mathbf{C}_n$ in which case $\gamma_1$ and $\gamma_2$ are the non-commutative analogues of its marginals.) We first prove a technical result.

**Lemma 1** *The function* $u(p, \gamma_{12}) \equiv \frac{1}{p-1}\Big(1 - \frac{\mathrm{Tr}_{12}\, \gamma_{12}^p}{\mathrm{Tr}_1\, \gamma_1^p}\Big)$ *is well-defined for* $p > 1$ *and* $\gamma_{12}$ *a density matrix. It can be extended by continuity to* $p \geq 1$ *and this extension satisfies*

$$u(1, \gamma_{12}) = -\frac{d}{dp} \frac{\mathrm{Tr}_{12}\, \gamma_{12}^p}{\mathrm{Tr}_1\, \gamma_1^p}\Big|_{p=1} = S(\gamma_{12}) - S(\gamma_1). \tag{2.6}$$

*Moreover, $u(p,\gamma_{12})$ is uniformly bounded in $\gamma_{12}$ for $p \in [1,2]$ and the continuity at $p=1$ is uniform in $\gamma_{12}$.*

**Proof:** It is well-known and straightforward to verify that, for any density matrix $\rho$ in $M_m$, $\lim_{p\to 1} \frac{1}{p-1}\big(1 - \operatorname{Tr}\rho^p\big) = S(\rho)$ and that $0 \le S(\rho) \le \log m$. It then follows that (2.6) holds; the convergence is uniform in $\gamma_{12}$ because the set of density matrices is compact. By the mean value theorem, for any fixed $p, \gamma_{12}$ one can find $\widetilde{p}$ with $1 \le \widetilde{p} \le p$ such that $u(p,\gamma_{12}) = -\frac{d}{dp} \frac{\operatorname{Tr}_{12}\gamma_{12}^{\widetilde{p}}}{\operatorname{Tr}_1 \gamma_1^{\widetilde{p}}}\Big|_{p=\widetilde{p}}$. Combining this with the fact that $\gamma \ge \gamma^{\widetilde{p}} \ge \gamma^2$ for any density matrix and $\widetilde{p} \in (1,2]$ gives the following bound

$$
\begin{aligned}
|u(p,\gamma_{12})| &= \left|\frac{\operatorname{Tr}_{12}\gamma_{12}^{\widetilde{p}} \log\gamma_{12} \operatorname{Tr}_1 \gamma_1^{\widetilde{p}} - \operatorname{Tr}_1 \gamma_1^{\widetilde{p}} \log\gamma_1 \operatorname{Tr}_{12}\gamma_{12}^{\widetilde{p}}}{\operatorname{Tr}_1 \gamma_1^{\widetilde{p}}}\right| && (2.7)\\
&\le \frac{S(\gamma_{12}) + S(\gamma_1)}{\operatorname{Tr}_1 \gamma_1^2} \qquad \textbf{QED} && (2.8)
\end{aligned}
$$

which is uniform in $p$ for $p \in (1,2]$.

The quantity $S_{\text{cond}}(\gamma_{12}) \equiv S(\gamma_{12}) - S(\gamma_1)$ is called the conditional entropy. Motivated by (2.6), we define the C.B. minimal entropy as

$$
S_{\text{CB,min}}(\Phi) = \inf_{\psi \in \mathbf{C}^d \otimes \mathbf{C}^d} S_{\text{cond}}\Big[(\mathcal{I}\otimes\Phi)\big(|\psi\rangle\langle\psi|\big)\Big] \tag{2.9}
$$

and observe that it satisfies the following.

**Theorem 2** *For any CPT map $\Phi$,*

$$
S_{\text{CB,min}}(\Phi) = \lim_{p\to 1+} \frac{1 - \big[\omega_p(\Phi)\big]^p}{p-1} \tag{2.10}
$$

*where $\omega_p(\Phi)$ is given by* (1.6).

**Proof:** With $\gamma_{12} = (\mathcal{I}\otimes\Phi)\big(|\psi\rangle\langle\psi|\big)$, one finds

$$
\begin{aligned}
S_{\text{CB,min}}(\Phi) &= \inf_{|\psi\rangle \in \mathbf{C}^d \otimes \mathbf{C}^d} u(1,\gamma_{12}) = \inf_{\psi} \lim_{p\to 1+} u(p,\gamma_{12}) \\
&= \inf_{\psi} \lim_{p\to 1+} \frac{1}{p-1}\left(1 - \frac{\operatorname{Tr}_{12}\gamma_{12}^p}{\operatorname{Tr}_1 \gamma_1^p}\right) \\
&= \lim_{p\to 1+} \inf_{\psi} \frac{1}{p-1}\left(1 - \frac{\operatorname{Tr}_{12}\gamma_{12}^p}{\operatorname{Tr}_1 \gamma_1^p}\right) && (2.11)\\
&= \lim_{p\to 1+} \frac{1}{p-1}\left(1 - \sup_{\psi} \frac{\operatorname{Tr}_{12}\gamma_{12}^p}{\operatorname{Tr}_1 \gamma_1^p}\right)
\end{aligned}
$$

where the interchange of $\lim_{p \to 1+}$ and $\inf_\psi$ is permitted by the uniformity in $\gamma_{12}$ of the continuity of $u(p, \gamma_{12})$ at $p = 1$.

**Theorem 3** *For all pairs of CPT maps $\Phi_A, \Phi_B$,*

$$S_{\mathrm{CB,min}}(\Phi_A \otimes \Phi_B) = S_{\mathrm{CB,min}}(\Phi_A) + S_{\mathrm{CB,min}}(\Phi_B)$$

**Proof:** The result follows easily from the observations above and (2.5).

$$\begin{aligned} S_{\mathrm{CB,min}}(\Phi_A \otimes \Phi_B) &= \lim_{p \to 1+} \frac{1 - \big[\omega_p(\Phi_A \otimes \Phi_B)\big]^p}{p-1} \\ &= \lim_{p \to 1+} \frac{1 - \big[\omega_p(\Phi_A)\big]^p \big[\omega(\Phi_B)\big]^p}{p-1} \\ &= \lim_{p \to 1+} \frac{1 - \big[\omega_p(\Phi_A)\big]^p}{p-1} + \left(\lim_{p \to 1+} \big[\omega_p(\Phi_A)\big]^p\right) \lim_{p \to 1+} \frac{1 - \big[\omega_p(\Phi_B)\big]^p}{p-1} \\ &= S_{\mathrm{CB,min}}(\Phi_A) + S_{\mathrm{CB,min}}(\Phi_B) \end{aligned} \tag{2.12}$$

where we used $\lim_{p \to 1+} \big[\omega_p(\Phi_A)\big]^p = 1$. **QED**

This result relies on (2.5) which is a special case of Theorem 11 with $q = 1$. Recently, Jencova [20] found a simple direct proof of (2.5).

## 2.3 Proof of CB additivity from SSA

Recall that any CPT map $\Phi$ can be represented in the form

$$\Phi(\rho) = \mathrm{Tr}_E U_{AE}\, \rho \otimes \tau_E\, U_{AE}^\dagger \tag{2.13}$$

with $U_{AE}$ unitary and $\tau_E$ a pure reference state on the environment. The following key result follows from standard purification arguments (which are summarized in Appendix A).

**Lemma 4** *Let the CPT map $\Phi$ have a representation as in* (2.13). *One can find a reference system $R$ and a pure state $|\psi_{RA}\rangle\langle\psi_{RA}|$ such that $\mathrm{Tr}_R |\psi_{RA}\rangle\langle\psi_{RA}| = \rho$. Define $\gamma_{REA} = (I_R \otimes U_{AE})\big(|\psi_{RA}\rangle\langle\psi_{RA}| \otimes \tau_E\big)(I_R \otimes U_{AE})^\dagger$. Then $\gamma_{REA}$ is also pure and*

$$S(\gamma_{EA}) - S(\gamma_E) = S(\gamma_R) - S(\gamma_{RA}) \tag{2.14}$$

*where the reduced density matrices are defined via partial traces.*

It follows from (1.8) that the conditional entropy is subadditive, i.e., for any state $\gamma_{E_1E_2A_1A_2}$,

$$S(\gamma_{E_1E_2A_1A_2}) - S(\gamma_{E_1E_2}) \quad\leq\quad S(\gamma_{E_1A_1}) - S(\gamma_{E_1}) + S(\gamma_{E_2A_2}) - S(\gamma_{E_2}) \tag{2.15}$$

This was proved by Nielsen [36] and appears as Theorem 11.16 in [37]. It follows easily from the observation that (2.15) is the sum of the following pair of inequalities, which are special cases of SSA

$$\begin{aligned} S(\gamma_{E_1E_2A_1A_2}) + S(\gamma_{E_1}) &\leq S(\gamma_{E_1A_1}) + S(\gamma_{E_1E_2A_2}) \\ S(\gamma_{E_1E_2A_2}) + S(\gamma_{E_2}) &\leq S(\gamma_{E_1E_2}) + S(\gamma_{E_2A_2}). \end{aligned}$$

Now define

$$S_{\mathrm{CB,min}}(\Phi) = \inf_{\psi} \Big( S\big[(\mathcal{I}\otimes\Phi)\big(|\psi\rangle\langle\psi|\big)\big] - S\big[\mathrm{Tr}_A|\psi\rangle\langle\psi|\big]\Big), \tag{2.16}$$

Let $\Psi_{RA_1A_2}$ denote the minimizer for $\Phi_1\otimes\Phi_2$ and

$$\gamma_{R_1R_2A_1A_2E_1E_2} = (I_R\otimes U_{A_1E_1A_2E_2})\big(|\psi_{RA_1A_2}\rangle\langle\psi_{RA_1A_2}|\otimes\tau_{E_1E_2}\big)(I_R\otimes U_{A_1E_1A_2E_2})^{\dagger}. \tag{2.17}$$

Then

$$\begin{aligned} S_{\mathrm{CB,min}}(\Phi_1\otimes\Phi_2) &= S(\gamma_{R_1R_2A_1A_2}) - S(\gamma_{R_1R_2}) \\ &= S(\gamma_{E_1E_2}) - S(\gamma_{E_1E_2A_1A_2}) \\ &\geq S(\gamma_{E_1}) - S(\gamma_{E_1A_1}) + S(\gamma_{E_2}) - S(\gamma_{E_2A_2}). \end{aligned} \tag{2.18}$$

Next, use the lemma to find purifications $\psi'_{RA}$ and $\psi''_{RA}$ so that the last line above

$$\begin{aligned} &= S(\gamma'_{R_1A_1}) - S(\gamma'_{R_1}) + S(\gamma''_{R_2A_2}) - S(\gamma''_{R_2}) \\ &\geq S_{\mathrm{CB,min}}(\Phi_1) + S_{\mathrm{CB,min}}(\Phi_2). \end{aligned} \tag{2.19}$$

The reverse inequality can be obtained using product $\Psi$.

# 3 Completely bounded norms

## 3.1 Definitions

For the applications in this paper, we can define the completely bounded (CB) norm of a map $\Phi : L_q(M_m) \mapsto L_p(M_n)$ as

$$\|\Phi\|_{\mathrm{CB},q\to p} \equiv \sup_d \|\mathcal{I}_d\otimes\Phi\|_{(\infty,q)\to(\infty,p)} = \sup_d \left( \sup_Y \frac{\|(\mathcal{I}_d\otimes\Phi)(Y)\|_{(\infty,p)}}{\|Y\|_{(\infty,q)}} \right). \tag{3.1}$$

with

$$\|Y\|_{(\infty,p)} \equiv \|Y\|_{L_\infty(M_d;L_p(M_n))} = \sup_{A,B\in M_d} \frac{\|(A\otimes \mathrm{I}_n)Y(B\otimes \mathrm{I}_n)\|_p}{\|A\|_{2p}\,\|B\|_{2p}}. \tag{3.2}$$

Effros and Ruan [12, 13] introduced the norm $\|Y\|_{(1,p)}$. Pisier [39, 40] subsequently used complex interpolation between them to define a norm $\|Y\|_{(t,p)}$ for any $1 < t < \infty$. He showed (Theorem 1.5 in [40]) that the norm obtained by this procedure satisfies

$$\|Y\|_{(t,p)} \equiv \|Y\|_{L_t(M_d;L_p(M_n))} = \inf_{\substack{Y=(A\otimes \mathrm{I}_n)Z(B\otimes \mathrm{I}_n)\\ A,B\in M_d}} \|A\|_{2t}\,\|B\|_{2t}\,\|Z\|_{(\infty,p)}, \tag{3.3}$$

which we can regard as its definition. The vector space $M_d \otimes M_n$ equipped with the norm (3.3) is a Banach space which we denote by $L_t(M_d;L_p(M_n))$. Given an operator $\Omega : L_t(M_d;L_q(M_m)) \mapsto L_s(M_{d'};L_p(M_n))$, the usual norm for linear maps from one Banach space to another becomes

$$\|\Omega\| \equiv \|\Omega\|_{(t,q)\to(s,p)} = \sup_{Q\in M_d\otimes M_m} \frac{\|\Omega(Q)\|_{(s,p)}}{\|Q\|_{(t,q)}}. \tag{3.4}$$

Theorem 1.5 and Lemma 1.7 in Pisier [40] show that one can use this norm to obtain another expression for the CB norm

$$\|\Phi\|_{\mathrm{CB},q\to p} \equiv \sup_d \|\mathcal{I}_d \otimes \Phi\|_{(t,q)\to(t,p)} = \sup_d \left( \sup_Y \frac{\|(\mathcal{I}_d\otimes\Phi)(Y)\|_{(t,p)}}{\|Y\|_{(t,q)}} \right) \tag{3.5}$$

valid for all $t \geq 1$. In effect, we can replace $\infty$ in (3.1) by any $t \geq 1$. In working with the CB norm, we will find it convenient to choose $t = q$ when $q \leq p$ and $t = p$ when $q \geq p$. Thus our working definition of the CB norm is (3.5) with $t = \min\{q,p\}$. For the applications considered in Sections 2 and 5, this becomes $t = q = 1$.

**Remark:** When $X > 0$, Hölder's inequality implies

$$\|AXB^\dagger\|_p \leq \sqrt{\|AXA^\dagger\|_p\,\|BXB^\dagger\|_p} \leq \max\{\|AXA^\dagger\|_p\,,\,\|BXB^\dagger\|_p\}$$

and the unitary invariance of the norm implies that $\|AXA^\dagger\|_p = \|\,|A|\,X\,|A|\,\|_p$. Therefore, when $X \geq 0$, we can replace any expression of the form $\sup_{A,B}\|AXB^\dagger\|_p$ by $\sup_{A>0}\|AXA\|_p$ irrespective of what other restrictions may be placed upon $A, B$. We will show that for CP maps, the CB norm is unchanged if the supremum is taken over $Y > 0$. (See Section 3.2, and Theorem 12 and Corollary 14 in Section 4.) Thus, when working with CP maps, one can generally assume that $A = B > 0$ in expressions for $\|Y\|_{(q,p)}$.

When $Y > 0$ combining (3.2) and (3.3) gives the identity,

$$\|Y\|_{(p,p)} = \inf_{\substack{B>0 \\ \operatorname{Tr} B=1}} \sup_{\substack{A>0 \\ \operatorname{Tr} A=1}} \|(A \otimes \mathrm{I}_n)^{\frac{1}{2p}}(B \otimes \mathrm{I}_n)^{-\frac{1}{2p}} Y (B \otimes \mathrm{I}_n)^{-\frac{1}{2p}}(A \otimes \mathrm{I}_n)^{\frac{1}{2p}}\|_p \tag{3.6}$$

for all $p \geq 1$. Since Theorem 7 implies that $\|Y\|_{(p,p)} = \|Y\|_p$, this gives a variational expression for the usual $p$-norm on $M_{dn} \simeq M_d \otimes M_n$. The choice $n = 1$ yields a max-min principle for the $p$-norm on $M_d$.

The Banach space $L_t(M_d; L_p(M_n))$ is a special case of a more general Banach space $L_t(M_d; E)$ for which a norm is defined on $d \times d$ matrices with entries in an operator space $E$ as described in Section 3.3. Because we use here only operators $\Phi : L_q(M_m) \mapsto L_p(M_n)$ rather than the general situation of operators $\Omega : E \mapsto F$ between Banach spaces $E, F$, we give explicit expressions only for norms on $L_t(M_d; L_p(M_n))$. On a few occasions we need to consider spaces $L_t(M_d; E)$ with $E = L_q(M_m; L_p(M_n))$; we denote the norm on these space by $\|Y\|_{(t,q,p)}$. In general we will only encounter triples with two distinct indices and will not need additional expressions for these norms. Such cases as $\|Y\|_{(q,q,p)}$ reduce to $L_q(M_{dm}; L_p(M_n))$ via the isomorphism between $M_{dm} \otimes M_n \simeq M_d \otimes M_m \otimes M_n$; most situations require only comparisons via Minkowski type inequalities given in Section 3.4. In section 3.5 we show that $\|Y\|_{(1,p,1)} = \|Y\|_{(1,p)}$; this is needed only for the application in Section 6.

## 3.2 An important lemma

We illustrate the use of (3.3) by proving the following lemma, which is a special case of a more general result in [23]. It plays a key role in the multiplicativity results of Section 4.2 for $q \geq p$. Although not needed for our main result, it also has important implications when $q \leq p$. We first define $\|\Phi\|^+_{q\to p} = \sup_{Q>0} \dfrac{\|\Phi(Q)\|_p}{\|Q\|_q}$.

**Lemma 5** *Let $\Phi : L_q(M_m) \mapsto L_p(M_n)$ be a CP map. Then for every $r \geq 1$ the map $\Phi \otimes \mathcal{I}_d : L_q(M_m; L_r(M_d)) \mapsto L_p(M_n; L_r(M_d))$ satisfies*

$$\|\Phi \otimes \mathcal{I}_d\|_{(q,r)\to(p,r)} \leq \|\Phi\|^+_{q\to p} \tag{3.7}$$

**Proof of Lemma:** For any $Q$ (3.3) implies that one can find $A, Y$ such that $Q = (A \otimes \mathrm{I}) Y (B \otimes \mathrm{I})$ and $\|Q\|_{(q,r)} = \|A\|_{2q} \|B\|_{2q} \|Y\|_{(\infty,r)}$. Since $\Phi$ is completely positive,

one can find $K_j$ satisfying (3.29). Let $V_A$ denote the block row vector with elements $\big(K_1 A \otimes \mathrm{I}_d,\, K_2 A \otimes \mathrm{I}_d,\, \ldots, K_m A \otimes \mathrm{I}_d\big)$, and similarly for $B$. Then

$$(\Phi \otimes \mathcal{I}_d)(Q) = V_A(\mathrm{I}_\nu \otimes Y)V_B^\dagger = \sum_j (K_j A \otimes \mathrm{I}_d) Y (B K_j^\dagger \otimes \mathrm{I}_d). \tag{3.8}$$

(Note that $\mathrm{I}_\nu \otimes Y$ denotes a block diagonal matrix with $Y$ along the diagonal with $\mathrm{I}_\nu$ the identity in an additional reference space used to implement the representation (3.29). $Y$ itself is in the tensor product space $M_m \otimes M_d$ on which $\Phi \otimes \mathcal{I}_d$ acts; $K_j$ and $A$ are in $M_m$. We can extend $V$ to an element of $M_\nu \otimes M_m \otimes M_d$ by adding rows of zero blocks; i.e., to $\sum_{i,j=1}^{\nu} \delta_{i1} |i\rangle\langle j| \otimes K_j A \otimes \mathrm{I}_d$.) Therefore, applying (3.3) on this extended space gives

$$\begin{aligned}
\|(\Phi \otimes \mathcal{I}_d)(Q)\|_{(p,r)} &\leq \|V_A^\dagger V_A\|_p^{1/2} \, \|V_B^\dagger V_B\|_p^{1/2} \|\mathrm{I}_\nu \otimes Y\|_{(\infty,\infty,r)} && (3.9)\\
&= \Big\| \sum_j K_j^\dagger A^\dagger A K_j \Big\|_p \, \Big\| \sum_j K_j^\dagger B^\dagger B K_j \Big\|_p^{1/2} \, \|Y\|_{(\infty,r)} \\
&= \|\Phi(A^\dagger A)\|_p^{1/2} \, \|\Phi(A^\dagger A)\|_p^{1/2} \, \|Y\|_{(\infty,r)} && (3.10)\\
&\leq \|\Phi\|_{q\to p}^+ \, \|A^\dagger A\|_q^{1/2} \, \|B^\dagger B\|_q^{1/2} \, \|Y\|_{(\infty,r)} \\
&= \|\Phi\|_{q\to p}^+ \, \|Q\|_{(q,r)}
\end{aligned}$$

where we used $\|A^\dagger A\|_q = \big(\|A\|_{2q}\big)^2$. **QED**

The following corollary implies that for any $p, q$, the norm $\|\Phi\|_{q\to p}$ is achieved on a positive semi-definite matrix $Q > 0$. This was proved earlier by Watrous [51], resolving a question raised in [26]. In Section 4, we will see that a similar result holds for CB norms of CP maps. This is stated as Theorem 12 for $q \leq p$ and Corollary 14 for $q \geq p$.

**Corollary 6** *Let $\Phi : L_q(M_m) \mapsto L_p(M_n)$ be a CP map. Then for all $q, p \geq 1$, the norm $\|\Phi\|_{q\to p} = \|\Phi\|_{q\to p}^+$*

**Proof:** The choice $d = 1$ in Lemma 5 gives $\|\Phi\|_{q\to p} \leq \|\Phi\|_{q\to p}^+$. Since the reverse inequality always holds, the result follows.

Note that one can similarly conclude that $\sup_d \|\Phi \otimes \mathcal{I}_d\|_{(q,t)\to(p,t)} = \|\Phi\|_{q\to p}^+$ so that nothing would be gained by defining an alternative to the CB norm in this way. In Section 5 we show that the depolarizing channel gives an explicit example of a map with $\|\Phi\|_{\mathrm{CB},1\to p} > \|\Phi\|_{1\to p}$. It is worth commenting on the difference between this result and the proof by Amosov, Holevo and Werner [3] that $\|\mathcal{I} \otimes \Phi\|_{(1,1)\to(p,p)} = \|\Phi\|_{1\to p}$. In the latter, the identity is viewed as an isometry from one Banach space $L_q(M_d)$ to another, $L_p(M_d)$. In the case of the CB norm, the identity is viewed as a map from the Banach

space $L_t(M_d)$ onto itself. Thus, we consider $\mathcal{I}_d \otimes \Phi$ with $\mathcal{I}_d : L_t(M_d) \to L_t(M_d)$ and $\Phi : L_q(M_m) \mapsto L_p(M_n)$, for which we need to consider what norm should be used on the domain $M_d \otimes M_m$ if $t \neq q$ or on the range $M_d \otimes M_n$ if $t \neq p$? When $q \neq p$ this question is unavoidable. One needs a norm which acts like $L_t$ on $M_d$ and $L_p$ on $M_n$, and (3.2) provides such a norm. Some of the motivation for the definitions used here is sketched in the next section.

For a discussion of the stability properties of $\|\mathcal{I} \otimes \Phi\|_{(q,q)\to(p,p)}$ see Kitaev [27] and Watrous [51]. Note that in the case $q = p$, the two types of norms for the extension $\Phi \otimes \mathcal{I}_d$ coincide and our results imply that for CP maps $\|\Phi\|_{\mathrm{CB,p\to p}} = \|\Phi \otimes \mathcal{I}_d\|_{(p,p)\to(p,p)} = \|\Phi\|^+_{p\to p}$. However, for measuring the difference between channels [27, 51], one is primarily interested in maps of the form $\Phi_1 - \Phi_2$ which are not CP.

### 3.3 Operator spaces

The Banach space $E = L_p(M_n)$ together with the sequence of norms on the spaces $L_\infty(M_d; L_p(M_n))$ with $d = 1, 2, \ldots$ form what is known as an *operator space*. More generally, an operator space is a Banach space $E$ and a sequence of norms defined on the spaces $M_d(E)$, whose elements are $d \times d$ matrices with elements in $E$, with certain properties that guarantee that $E$ can be embedded in $\mathcal{B}(\mathcal{H})$, the bounded operators on some Hilbert space $\mathcal{H}$. Alternatively, one can begin with a subspace $E \subset \mathcal{B}(\mathcal{H})$; then the norm in $M_d(E)$ is given by the inclusion $M_d(E) \subset M_d\big(\mathcal{B}(\mathcal{H})\big) \simeq \mathcal{B}(\mathcal{H}^{\otimes d})$ consistent with interpreting an element of $M_d(E)$ as a block matrix. (Usually such a situation is considered a concrete operator space in contrast to an abstract operator space given by matrix norms satisfying Ruan's axioms [13, 41, 44].) The only operator spaces we use in this paper are those with $E = L_p(M_n)$ and, occasionally, $E = L_t(M_d; L_p(M_n))$. Although a concrete representation for even these spaces is not known, the explicit expressions for the norms given in Sections 3.1 and 3.5 suffice for many purposes. (The reader who wishes to explore the literature should be aware that most of it is written in terms of $L_t(M_d; E)$ rather than $L_t(M_d; L_p(M_n))$ and that the notation $S_t(M_d; E)$ (for Schatten norm) is more common than $L_t$.)

For maps from $\mathcal{B}(\mathcal{H})$ to $\mathcal{B}(\mathcal{K})$ complete boundedness is just uniform boundedness for the sequence of norms of $\mathcal{I}_d \otimes \Phi$ . This notion is built in a manner analogous to the familiar notion of complete positivity. In a similar way, one can define other "complete" notions, such as complete isometry based on the behavior of $\mathcal{I}_d \otimes \Phi$.

The particular type of operator space considered here is called a "vector-valued $L_p$ space". We have already remarked on the need to define a norm on $L_t(M_d; L_p(M_n))$ to

give a non-commutative generalization of the classical Banach space $\ell_t(\ell_p)$. Unfortunately, such naive generalizations as $\big(\sum_{jk} \|Y_{jk}\|_p^t\big)^{1/t}$ or $\big(\mathrm{Tr}_1\big(\mathrm{Tr}_2|Y|^p\big)^{t/p}\big)^{1/t}$ do not even define norms. The norms described in Section 3.1, although difficult to work with, yield an elegant structure with the following properties.

a) for the subalgebra of diagonal matrices the norm on $L_t(M_d; L_p(M_n))$ reduces to that on $\ell_t(\ell_q)$.

b) When $Y = A \otimes B$ is a tensor product, $\|Y\|_{t,p} = \|A\|_t \, \|B\|_p = \big(\mathrm{Tr}\,|A|^t\big)^{1/t} \big(\mathrm{Tr}\,|B|^p\big)^{1/p}$.

c) The Banach space duality between $L_p$ and $L_{p'}$ with $\frac{1}{p} + \frac{1}{p'} = 1$ generalizes to

$$L_q(M_d; L_p(M_n))^* = L_{q'}(M_d; L_{p'}(M_n)). \tag{3.11}$$

d) The collection of norms on $\{L_t(M_d; L_p(M_n))\}$ can be obtained from some (abstract) embedding of $L_p(M_d)$ into $\mathcal{B}(\mathcal{H})$ providing the operator space structure of $L_p(M_d)$.

e) The structure of $L_t(M_d; L_p(M_n))$ can be used to develop a theory of vector-valued non-commutative integration which generalizes the theory of non-commutative integration developed by Segal [45] and Nelson [35].

Although not used explicitly, properties (c) and (e) play an important role in our results. Consequences of (e) described in Section 3.4 play a key role in the proofs in Section 4 and Section 6. Theorem 10, which gives the simple expression (1.6) for the CB norm in the case $1 \to p$, is an immediate consequence of a fundamental duality theorem.

For general information on operator spaces, see Paulsen [38], Effros and Ruan [13] or Pisier [41]. The theory of non-commutative vector valued $L_p$ spaces was developed by Pisier in two monographs [39] and [40]. Additional developments can be found in [21] and [41].

## 3.4 Fubini and Minkowski generalizations

Because vector valued $L_p$-spaces permit the development of a consistent theory of vector-valued non-commutative integration, one would expect generalizations of fundamental integration theorems. This is indeed the case, and analogues of both Fubini's theorem and Minkowski's inequality play an important role in the results that follow.

First, Theorem 1.9 in [40] gives a non-commutative version of Fubini's theorem.

**Theorem 7** *For any $1 \le p \le \infty$, the isomorphisms $L_p(M_d; L_p(M_n)) \simeq L_p(M_d \otimes M_n) \simeq L_p(M_{dn})$ hold in the sense of complete isometry, which implies that for all $W \in M_d \otimes M_n$,*

$$\|W\|_{L_p(M_d;L_p(M_n))} = \|W\|_{L_p(M_n;L_p(M_d))} = \|W\|_p = \big(\mathrm{Tr}\, W^p\big)^{1/p}. \tag{3.12}$$

The next result, which is Theorem 1.10 in [40], will lead to non-commutative versions of Minkowski's inequality and deals with the flip map $F$ which takes $A \otimes B \mapsto B \otimes A$ and is then extended by linearity to arbitrary elements of a tensor product space so that $W_{12} \mapsto W_{21}$.

**Theorem 8** *For $q \le p$, the flip map $F : L_q(M_d; L_p(M_n)) \mapsto L_p(M_n; L_q(M_d))$ is a complete contraction.*

The fact that $F$ is a contraction yields an analogue of Minkowski's inequality for matrices.

$$\|W_{21}\|_{(p,q)} = \|F(W_{12})\|_{(p,q)} \le \|W_{12}\|_{(q,p)} \qquad \text{for} \quad q \le p. \tag{3.13}$$

The fact that $F$ is a *complete* contraction means that $\mathcal{I} \otimes F$ is also a contraction which yields a triple Minkowski inequality

$$\|W_{132}\|_{(q,p,q)} \le \|W_{123}\|_{(q,q,p)} \tag{3.14}$$

when $q \le p$.

**Remark:** To see why we regard (3.13) as a non-commutative version of Minkowski's inequality, recall the usual $\ell_p(\ell_q)$ version. For $t \ge 1$, $\Big[\sum_j \big(\sum_k |a_{jk}|\big)^t\Big]^{1/t} \le \sum_k \Big(\sum_j |a_{jk}|^t\Big)^{1/t}$, and Carlen and Lieb [8] extended this to positive semi-definite matrices

$$\Big[\mathrm{Tr}_1\Big(\mathrm{Tr}_2\, Q_{12}\Big)^t\Big]^{1/t} \le \mathrm{Tr}_2\Big(\mathrm{Tr}_1\, Q_{12}^t\Big)^{1/t} \tag{3.15}$$

As in the case of the classical inequalities, (3.15) holds for $t \ge 1$ and the reverse inequality holds for $t \le 1$. Moreover, it follows that for $R \ge 0$

$$\Big[\mathrm{Tr}_1\Big(\mathrm{Tr}_2\, R_{12}^q\Big)^{p/q}\Big]^{1/p} \le \Big[\mathrm{Tr}_2\Big(\mathrm{Tr}_1\, R_{12}^p\Big)^{q/p}\Big]^{1/q} \qquad \text{for} \quad q \le p. \tag{3.16}$$

To see that (3.16) and (3.15) are equivalent, let $t = p/q$, and $Q_{12} = R_{12}^p$. Then raising both sides of (3.16) to the $q$-th power yields (3.15).

In general, the quantity $\left[\operatorname{Tr}_1\left(\operatorname{Tr}_2 R^p\right)^{q/p}\right]^{1/q}$ does not define a norm. Carlen and Lieb [8] conjectured that $\operatorname{Tr}_1\left(\operatorname{Tr}_2 R^p\right)^{1/p}$ does define a norm for $1 \leq p \leq 2$, but proved it only in the case $p = 2$. (For $p > 2$ it can be shown not to be a norm.) Their conjecture is that

$$\operatorname{Tr}_3\left[\operatorname{Tr}_2\left(\operatorname{Tr}_1 Q_{123}\right)^t\right]^{1/t} \leq \operatorname{Tr}_{1,3}\left(\operatorname{Tr}_2 Q_{123}^t\right)^{1/t} \tag{3.17}$$

which is very similar in form to (3.14) with $q = 1, p = t$.

## 3.5 More facts about $L_q(M_d; L_p(M_n))$ norms

We now state two additional formulas for norms on $L_q(M_d; L_p(M_n))$. Although not needed for the main result, some consequences are needed for Theorem 12 and in Section 6. For detailed proofs see [21].

We state both under the assumption $1 \leq q \leq p \leq \infty$ and $\frac{1}{q} = \frac{1}{p} + \frac{1}{r}$. Then

$$\|Y\|_{(p,q)} \equiv \|Y\|_{L_p(M_d;L_q(M_n))} = \sup_{A,B \in M_d} \frac{\|(A \otimes \mathrm{I}_n) Y (B \otimes \mathrm{I}_n)\|_q}{\|A\|_{2r}\,\|B\|_{2r}} \tag{3.18}$$

and

$$\|Y\|_{(q,p)} \equiv \|Y\|_{L_q(M_d;L_p(M_n))} = \inf_{\substack{Y = (A \otimes \mathrm{I}_n) Z (B \otimes \mathrm{I}_n) \\ A,B \in M_d}} \|A\|_{2r}\,\|B\|_{2r}\,\|Z\|_p \tag{3.19}$$

Moreover, when $Y > 0$ is positive semi-definite, one can restrict both optimizations to $A = B > 0$. In the case $X > 0$, $q = 1$, (3.18) becomes

$$\begin{aligned} \|X_{12}\|_{(p,1)} &= \sup_{A>0} \frac{\|(A \otimes \mathrm{I}_n) X_{12} (A \otimes \mathrm{I}_n)\|_1}{\|A\|_{2p'}^2} \\ &= \sup_{A>0} \frac{\operatorname{Tr} A^2 X_1}{\|A^2\|_{p'}} = \|X_1\|_p \end{aligned} \tag{3.20}$$

and (3.19) can be rewritten as

$$\|X\|_{(1,p)} = \inf_{\substack{A>0 \\ X = (A \otimes \mathrm{I}_n) Z (A \otimes \mathrm{I}_n)}} \|A\|_{2p'}^2\,\|Z\|_p \tag{3.21}$$

$$= \inf_{B>0,\,\|B\|_1 = 1} \|(B^{-1/2p'} \otimes \mathrm{I}_n)\, X\, (B^{-1/2p'} \otimes \mathrm{I}_n)\|_p$$

$$= \inf_{B>0,\,\|B\|_1 = 1} \|(B^{-\frac{1}{2}(1-\frac{1}{p})} \otimes \mathrm{I}_n)\, X\, (B^{-\frac{1}{2}(1-\frac{1}{p})} \otimes \mathrm{I}_n)\|_p \tag{3.22}$$

In Section 6, we will also need

$$\begin{aligned}
\|W_{132}\|_{(1,p,1)} &= \|W_{132}\|_{L_1(M_d;L_p(M_n;L_1(M_m)))} \\
&= \inf_{\substack{A\in M_d, A>0 \\ W_{132}=(A\otimes I_{32})Z_{132}(A\otimes I_{32})}} \|A\|_{2p'}^2 \, \|Z_{132}\|_{(p,p,1)} \qquad (3.23)\\
&= \inf_{B_1>0, \|B_1\|_1} \|(B_1^{-1/2p'}\otimes I_3 \otimes I_2)W_{132}(B_1^{-1/2p'}\otimes I_3\otimes I_2)\|_{(p,p,1)} \\
&= \inf_{B_1>0, \|B_1\|_1} \|(B_1^{-1/2p'}\otimes I_3)W_{13}(B_1^{-1/2p'}\otimes I_3)\|_{(p,p)} \qquad (3.24)\\
&= \|W_{13}\|_{(1,p)}
\end{aligned}$$

where (3.23) is proved in [21] and the reductions which follow used (3.20) and (3.18).

**Lemma 9** *When $1 \le q \le p \le \infty$ and $X$ is a contraction, then*

$$\|C^\dagger X D\|_{(q,p)} \le \big(\|C^\dagger C\|_{(q,p)} \; \|D^\dagger D\|_{(q,p)}\big)^{1/2} \qquad (3.25)$$

**Proof:** It follows from (3.19) that one can find $A, B \in M_d$ and $Y, Z \in M_{dn}$ such that $A, B > 0$, $\|A\|_{2r} = \|B\|_{2r} = 1$, $Y, Z > 0$ and

$$\begin{aligned}
C^\dagger C &= (A\otimes I_n)Y(A\otimes I_n) \qquad & \|C^\dagger C\|_{(q,p)} &= \|(A\otimes I_n)Y(A\otimes I_n)\|_p \\
D^\dagger D &= (B\otimes I_n)Z(B\otimes I_n) \qquad & \|D^\dagger D\|_{(q,p)} &= \|(B\otimes I_n)Z(B\otimes I_n)\|_p \, .
\end{aligned}$$

Moreover, there are partial isometries, $V, W$ such that $C = VY^{1/2}(A\otimes I_n)$ and $D = WZ^{1/2}(B\otimes I_n)$. Then

$$C^\dagger X D = (A\otimes I_n)Y^{1/2}V^\dagger X W Z^{1/2}(B\otimes I_n) \qquad (3.26)$$

and it follows from (3.19) and Hölder's inequality that

$$\begin{aligned}
\|C^\dagger X D\|_{(q,p)} &\le \|(A\otimes I_n)Y^{1/2}\, V^\dagger X W\, Z^{1/2}(B\otimes I_n)\|_p \qquad (3.27)\\
&\le \|(A\otimes I_n)Y A\otimes I_n)\|_p^{1/2} \, \|V^\dagger X W\|_\infty \, \|(B\otimes I_n)Z B\otimes I_n)\|_p^{1/2} \\
&= \|C^\dagger C\|_{(q,p)}^{1/2} \; \|D^\dagger D\|_{(q,p)}^{1/2} \qquad \textbf{QED}
\end{aligned}$$

## 3.6 State representative of a map

A linear map $\Phi : M_d \mapsto M_d$ can be associated with a block matrix in which the $j,k$ block is the matrix $\Phi\big(|e_j\rangle\langle e_k|\big)$ in the standard basis. This is often called the "Choi-Jamiolkowski

matrix" or "state representative" in quantum information theory and will be denoted $X_\Phi$. Thus,

$$X_\Phi \;=\; \sum_{jk} |e_j\rangle\langle e_k| \otimes \Phi\big(|e_j\rangle\langle e_k|\big) \tag{3.28}$$

Choi [9] showed that the map $\Phi$ is CP if and only if $X_\Phi$ is positive semi-definite. Conversely given a (positive semi-definite) $d^2 \times d^2$ matrix $X$, one can use (3.28) to define a CP map $\Phi$. In addition, Choi showed that the eigenvectors of $X_\Phi$ can be rearranged to yield operators, $K_j$ such that

$$\Phi(Q) = \sum_j K_j Q K_j^\dagger. \tag{3.29}$$

This result representation was obtained independently by Kraus [29, 30] and can be recovered from that of Stinespring [50].

For every CP map $\Phi$ with Choi matrix $X_\Phi$, it follows from (3.28) that

$$\begin{aligned} \|(A \otimes \mathrm{I}) X_\Phi (A \otimes \mathrm{I})\|_p \;&=\; \|\sum_{jk} A|e_j\rangle\langle e_k|A \otimes \Phi\big(|e_j\rangle\langle e_k|\big)\|_p \qquad (3.30)\\ &=\; \|(\mathcal{I} \otimes \Phi)\big(|\psi_A\rangle\langle\psi_A|\big)\|_p \end{aligned}$$

where the last equality follows if we choose $|\psi_A\rangle = \sum_j A|e_j\rangle \otimes |e_j\rangle$

**Theorem 10** *For any CP map $\Phi$,*

$$\|\Phi\|_{\mathrm{CB},1\to p} = \|X_\Phi\|_{(\infty,p)} = \sup_{\|\psi\|=1} \frac{\|(\mathcal{I} \otimes \Phi)\big(|\psi\rangle\langle\psi|\big)\|_p}{\|\mathrm{Tr}_2\,(|\psi\rangle\langle\psi|)\|_p} \;\equiv\; \omega_p(\Phi) \tag{3.31}$$

**Proof:** This result requires a fundamental duality result proved by Blecher and Paulsen [7] and by Effros and Ruan [12, 13] and described in Section 2.3 of [41]. It states that

$$\|\Phi\|_{\mathrm{CB},1\to p} = \|\Phi^*\|_{\mathrm{CB},p'\to\infty} = \|X_\Phi\|_{(\infty,p)} \tag{3.32}$$

Using (3.2) gives

$$\begin{aligned} \|\Phi\|_{\mathrm{CB},1\to p} \;&=\; \sup_{A>0} \frac{\|(A \otimes \mathrm{I}) X_\Phi (A \otimes \mathrm{I})\|_p}{\|A^2\|_p} \\ &=\; \sup_{\psi} \frac{\|(\mathcal{I} \otimes \Phi)\big(|\psi\rangle\langle\psi|\big)\|_p}{\|\mathrm{Tr}_2\,(|\psi\rangle\langle\psi|)\|_p}. \end{aligned}$$

Since the ratio is unchanged if $|\psi\rangle$ is multiplied by a constant, one can restrict the supremum above to $\|\psi\| = 1$. **QED**

# 4 Multiplicativity for CB norms

## 4.1 $1 \leq q \leq p$

We now prove multiplicativity of the CB norm for maps $\Phi : L_q(M_m) \mapsto L_p(M_m)$ with $q \leq p$.

**Theorem 11** *Let $q \leq p$ and $\Phi_A : L_q(M_{m_A}) \mapsto L_p(M_{n_A})$ and $\Phi_B : L_q(M_{m_B}) \mapsto L_p(M_{n_B})$ be CP and CB. Then*

$$\|\Phi_A \otimes \Phi_B\|_{\mathrm{CB},q\to p} = \|\Phi_A\|_{\mathrm{CB},q\to p}\|\Phi_B\|_{\mathrm{CB},q\to p} \ . \tag{4.1}$$

**Proof:** Let $Q_{CAB}$ be in $M_d \otimes M_{m_A} \otimes M_{m_B}$ and $R_{CAB} = (\mathcal{I}_d \otimes \mathcal{I}_{m_A} \otimes \Phi_B)(Q_{CAB})$. Then using (3.14), one finds

$$\|\Phi_A \otimes \Phi_B\|_{\mathrm{CB},q\to p} = \sup_d \sup_{Q_{CAB}} \frac{\|(\mathcal{I}_d \otimes \Phi_A \otimes \Phi_B) Q_{CAB}\|_{(q,p,p)}}{\|Q_{CAB}\|_{(q,q,q)}} \tag{4.2}$$

$$\begin{aligned}
&= \sup_{Q_{CAB}} \frac{\|(\mathcal{I}_d \otimes \Phi_A \otimes \mathcal{I}_{n_B}) R_{CAB}\|_{(q,p,p)}}{\|R_{CAB}\|_{(q,q,p)}} \, \frac{\|(\mathcal{I}_d \otimes \mathcal{I}_{m_A} \otimes \Phi_B)(Q_{CAB})\|_{(q,q,p)}}{\|Q_{CAB}\|_{(q,q,q)}} \\
&\leq \sup_{R_{CBA}} \frac{\|(\mathcal{I}_d \otimes \mathcal{I}_{n_B} \otimes \Phi_A) R_{CBA}\|_{(q,p,p)}}{\|R_{CBA}\|_{(q,p,q)}} \, \frac{\|R_{CBA}\|_{(q,p,q)}}{\|R_{CAB}\|_{(q,q,p)}} \qquad (4.3) \\
&\qquad\qquad \times \sup_{Q_{CAB}} \frac{\|(\mathcal{I}_d \otimes \mathcal{I}_{m_A} \otimes \Phi_B)(Q_{CAB})\|_{(q,q,p)}}{\|Q_{CAB}\|_{(q,q,q)}} \\
&\leq \|\mathcal{I}_{n_B} \otimes \Phi_A\|_{\mathrm{CB},(p,q)\to(p,p)} \ \|\Phi_B\|_{\mathrm{CB},q\to p} \qquad (4.4) \\
&= \|\Phi_A\|_{\mathrm{CB},q\to p} \ \|\Phi_B\|_{\mathrm{CB},q\to p}.
\end{aligned}$$

For the last two lines, we used $\|\mathcal{I}_n \otimes \Phi_A\|_{\mathrm{CB},(p,q)\to(p,p)}$ to denote the CB norm of $\mathcal{I}_n \otimes \Phi_A : L_p(M_n; L_q(M_m)) \mapsto L_p(M_n; L_p(M_m))$ and then applied Corollary 1.2 in [40], which states that this is the same as the CB norm of $\Phi : L_q(M_m) \mapsto L_p(M_m)$.

To prove the reverse direction, we need a slight modification of the standard strategy of showing that the bound can be achieved with a tensor product. It can happen that the CB norm itself is not attained for any finite $I_d \otimes \Phi$ norm. Therefore, we first show that any finite product can be achieved, and then use the fact that the CB norm can be approximated arbitrarily closely by such a product.

Thus, we begin with the observation that for any $d$ and $X, Y$ in the unit balls for $L_q(M_d \otimes M_m)$ and $L_q(M_d \otimes M_n)$, there exist $Q, R > 0$ in the unit ball of $L_{2q}(M_d)$ such

that

$$\|(Q \otimes 1_m)[\mathcal{I}_d \otimes \Phi_A(X)](Q \otimes \mathrm{I}_m)\|_q = \|(\mathcal{I}_d \otimes \Phi_A(X))\|_{L_q(M_d;L_p(M_m))} \tag{4.5}$$

and

$$\|(R \otimes \mathrm{I}_n)(\mathcal{I}_d \otimes \Phi_B(Y))(R \otimes \mathrm{I}_n)\|_q = \|[\mathcal{I} \otimes \Phi_B(Y)]\|_{L_q(M_d;L_p(M_n))} \; . \tag{4.6}$$

Then, using Theorem 7, one finds

$$\begin{aligned} \|\Phi_A \otimes \Phi_B\|_{\mathrm{CB},q\to p} &\geq \|[\mathcal{I}_{M_{d^2}} \otimes (\Phi_A \otimes \Phi_B)](X \otimes Y)\|_{L_q(M_{d^2};L_p(M_{mn}))} \\ &\geq \|(Q \otimes R \otimes \mathrm{I}_{mn})[\mathcal{I}_{d^2} \otimes (\Phi_A \otimes \Phi_B)](X \otimes Y)(Q \otimes R \otimes \mathrm{I}_{mn})\|_q \\ &= \|(Q \otimes \mathrm{I})[\mathcal{I}_{M_d} \otimes \Phi_A(X)](Q \otimes \mathrm{I})\|_q \; \|(R \otimes \mathrm{I})[\mathcal{I}_{M_d} \otimes \Phi_B(Y)](R \otimes \mathrm{I})\|_q \\ &= \|(\mathcal{I}_{M_d} \otimes \Phi_A(X))\|_{L_q(M_d;L_p(M_m))} \; \|(\mathcal{I}_{M_d} \otimes \Phi_B(X))\|_{L_q(M_d;L_p(M_n))} \end{aligned} \tag{4.7}$$

Given $\epsilon > 0$, one can find $d, X, Y$ such that $\|\Phi_A\|_{\mathrm{CB},q\to p} < \epsilon + \|(\mathcal{I}_{M_d} \otimes \Phi_A(X))\|_{L_q(M_d;L_p(M_m))}$ and $\|\Phi_B\|_{\mathrm{CB},q\to p} < \epsilon + \|(\mathcal{I}_{M_d} \otimes \Phi_B(X))\|_{L_q(M_d;L_p(M_n))}$. Inserting this in (4.7) above gives

$$\|\Phi_A \otimes \Phi_B\|_{\mathrm{CB},q\to p} \;\geq\; \|\Phi_A\|_{\mathrm{CB},q\to p}\|\Phi_B\|_{\mathrm{CB},q\to p} \;-\; \epsilon\big(\|\Phi_A\|_{\mathrm{CB},q\to p} + \|\Phi_B\|_{\mathrm{CB},q\to p}\big) + O(\epsilon^2)$$

Since $\epsilon > 0$ is arbitrary, we can conclude that

$$\|\Phi_A \otimes \Phi_B\|_{\mathrm{CB},q\to p} \geq \|\Phi_A\|_{\mathrm{CB},q\to p}\|\Phi_B\|_{\mathrm{CB},q\to p}. \qquad \textbf{QED}$$

The next result implies that for CP maps, it suffices to restrict the supremum in the CB norm to positive semi-definite matrices.

**Theorem 12** *When $q \leq p$ and $\Phi : L_q(M_m) \mapsto L_p(M_n)$ is CP, $\|\mathcal{I}_d \otimes \Phi\|_{(q,q)\to(q,p)}$ is achieved with a positive semi-definite matrix, i.e., $\|\mathcal{I}_d \otimes \Phi\|_{(q,q)\to(q,p)} = \|\mathcal{I}_d \otimes \Phi\|^+_{(q,q)\to(q,p)}$.*

**Proof:** First use the polar decomposition of $Q \in M_{dm}$ to write $Q = Q_1^\dagger Q_2$ with $Q_1 = |Q|^{1/2}U,\; Q_2 = |Q|^{1/2}$ where $U$ is a partial isometry and $|Q| = (Q^\dagger Q)^{1/2}$. The matrix

$$\begin{pmatrix} Q_1^\dagger \\ Q_2^\dagger \end{pmatrix} \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} = \begin{pmatrix} Q_1^\dagger Q_1 & Q_1^\dagger Q_2 \\ Q_2^\dagger Q_1 & Q_2^\dagger Q_2 \end{pmatrix} = \begin{pmatrix} U^\dagger |Q| U & Q \\ |Q| & Q^\dagger \end{pmatrix} > 0 \tag{4.8}$$

is positive semi-definite. Since $\Phi$ is CP, so is $\mathcal{I} \otimes \Phi$ which implies that

$$\begin{pmatrix} (\mathcal{I} \otimes \Phi)(U^\dagger |Q| U) & (\mathcal{I} \otimes \Phi)(Q) \\ (\mathcal{I} \otimes \Phi)(Q^\dagger) & (\mathcal{I} \otimes \Phi)(|Q|) \end{pmatrix} > 0 \tag{4.9}$$

is positive semi-definite. We now use the fact that a $2\times 2$ block matrix $\begin{pmatrix} A & C \\ C^\dagger & B \end{pmatrix}$ with $A, B > 0$ is positive semi-definite if and only if $C = A^{1/2} X B^{1/2}$ with $X$ a contraction. Applying this to (4.9) gives

$$(\mathcal{I}\otimes\Phi)(Q) = [(\mathcal{I}\otimes\Phi)(U^\dagger|Q|U)]^{1/2}\, X\, [(\mathcal{I}\otimes\Phi)(|Q|)]^{1/2} \tag{4.10}$$

with $X$ a contraction. Therefore, it follows from (3.25) that

$$\begin{aligned}
\|(\mathcal{I}\otimes\Phi)(Q)\|_{(q,p)} &= \|[(\mathcal{I}\otimes\Phi)(U^\dagger|Q|U)]^{1/2}\, X\, [(\mathcal{I}\otimes\Phi)(|Q|)]^{1/2}\|_{(q,p)} \qquad (4.11)\\
&\leq \Big(\|(\mathcal{I}\otimes\Phi)(U^\dagger|Q|U)\|_{(q,p)}\, \|(\mathcal{I}\otimes\Phi)(|Q|)\|_{(q,p)}\Big)^{1/2}\\
&\leq \|\mathcal{I}\otimes\Phi\|^{+}_{(q,q)\to(q,p)} \left(\||Q|\|_q\, \|U^\dagger|Q|U\|_q\right)^{1/2}\\
&= \|\mathcal{I}\otimes\Phi\|^{+}_{(q,q)\to(q,p)}\, \||Q|\|_q \qquad \textbf{QED}
\end{aligned}$$

## 4.2 $q \geq p$

**Theorem 13** *Let $q \geq p$ and $\Phi_A : L_q(M_{m_A}) \to L_p(M_{n_A})$, $\Phi_B : L_q(M_{m_B}) \to L_p(M_{n_B})$ be maps which are both CP. Then*

*a)* $$\|\Phi\|_{\mathrm{CB},q\to p} = \|\Phi\|_{q\to p} = \|\Phi\|^{+}_{q\to p} \tag{4.12}$$

*b)* $$\|\Phi_A\otimes\Phi_B\|_{q\to p} = \|\Phi_A\|_{q\to p}\|\Phi_B\|_{q\to p} \tag{4.13}$$

*c)* $$\|\Phi_A\otimes\Phi_B\|_{\mathrm{CB},q\to p} = \|\Phi_A\|_{\mathrm{CB},q\to p}\|\Phi_B\|_{\mathrm{CB},q\to p} \quad . \tag{4.14}$$

Combining part (a) with Corollary 6 implies that it suffices to restrict the supremum in the CB norm to positive semi-definite matrices.

**Corollary 14** *When $q \geq p$ and $\Phi : L_q(M_m) \mapsto L_p(M_n)$ is CP, $\|\mathcal{I}_d\otimes\Phi\|_{\mathrm{CB},q\to p}$ is achieved with a positive semi-definite matrix.*

**Proof of Theorem 13:** To prove part (a), observe that

$$\begin{aligned}
\|\Phi\|_{\mathrm{CB},q\to p} &= \sup_d \left( \sup_{W_{AB}\in M_d\otimes M_m} \frac{\|(\mathcal{I}_d\otimes\Phi)(W_{AB})\|_p}{\|W_{AB}\|_{(p,q)}} \right)\\
&= \sup_d \left( \sup_{W_{AB}} \frac{\|(\mathcal{I}_d\otimes\Phi)(W_{AB})\|_p}{\|W_{BA}\|_{(q,p)}}\; \frac{\|W_{BA}\|_{(q,p)}}{\|W_{AB}\|_{(p,q)}} \right) \qquad (4.15)\\
&\leq \sup_d \sup_{W_{BA}\in M_m\otimes M_d} \frac{\|(\Phi\otimes\mathcal{I}_d)(W_{BA})\|_p}{\|W_{BA}\|_{(q,p)}}\\
&\leq \|\Phi\|^{+}_{q\to p} \qquad (4.16)
\end{aligned}$$

The first inequality follows from the fact that the second ratio in (4.15) is $\leq 1$ by (3.13) and the last inequality then follows from (3.7). When $d = 1$, the supremum over $W$ of the ratio in (4.15) is precisely $\|\Phi\|_{q\to p}$ which implies $\|\Phi\|_{\mathrm{CB},q\to p} \geq \|\Phi\|_{q\to p}$. This proves part (a).

To prove part (b), write $\Phi_A \otimes \Phi_B = (\Phi_A \otimes \mathcal{I})(\mathcal{I} \otimes \Phi_B)$ and for any $Q_{AB} \in M_{m_A} \otimes M_{m_B}$, let $R_{AB} = (\mathcal{I} \otimes \Phi_B)(Q)$. Then

$$\|\Phi_A \otimes \Phi_B\|_{q\to p} = \sup_Q \frac{\|(\Phi_A \otimes \Phi_B)(Q)\|_p}{\|Q\|_q} \tag{4.17}$$

$$\leq \sup_Q \frac{\|(\Phi_A \otimes \mathcal{I})(R_{AB})\|_p}{\|R_{AB}\|_{(q,p)}} \frac{\|R_{AB}\|_{(q,p)}}{\|R_{BA}\|_{(p,q)}} \frac{\|(\Phi_B \otimes \mathcal{I})(Q_{BA})\|_{(p,q)}}{\|Q\|_q}$$

$$\leq \sup_R \frac{\|(\Phi_A \otimes \mathcal{I})(R)\|_{(p,p)}}{\|R\|_{(q,p)}} \sup_{Q_{BA}} \frac{\|(\Phi_B \otimes \mathcal{I})(Q_{BA})\|_{(p,q)}}{\|Q_{BA}\|_{q,q}} \tag{4.18}$$

$$\leq \|\Phi_A\|_{q\to p} \, \|\Phi_B\|_{q\to p}$$

where we used (3.13), Fubini, and $R_{BA} = (\Phi_B \otimes \mathcal{I})(Q_{BA})$. This proves (b).

Part (c) then follows immediately from (a) and (b). **QED**

# 5 Applications of CB entropy

## 5.1 Examples and bounds

It is well-known that conditional information can be negative as well as positive. Therefore, it is not surprising that (1.1) can also be either positive or negative, depending on the channel $\Phi$. As in Section 1, we adopt the convention that $\gamma_{12} = (\mathcal{I} \otimes \Phi)\big(|\psi\rangle\langle\psi|\big)$. One has the general bounds

$$-S(\gamma_1) \leq S_{\mathrm{CB,min}}(\Phi) \leq S(\gamma_1) \tag{5.1}$$

which imply

$$-\log d \;\; \leq S_{\mathrm{CB,min}}(\Phi) \leq \log d. \tag{5.2}$$

The lower bound in (5.2) follows from the definition (1.3) and the positivity of the entropy $S(\gamma_{12}) > 0$; the upper bound follows from subadditivity $S(\gamma_{12}) \leq S(\gamma_1) + S(\gamma_2)$. The upper bound is attained if and only if the output $(\mathcal{I} \otimes \Phi)\big(|\psi\rangle\langle\psi|\big)$ is always a product.

The lower bound in (5.2) is attained for the identity channel, and the upper bound for the completely noisy channel $\Phi(\rho) = (\operatorname{Tr}\rho)\frac{1}{d}I$.

Next, consider the depolarizing channel $\Omega_\mu(\rho) = \mu\rho + (1-\mu)(\operatorname{Tr}\rho)\frac{1}{d}I$. This channel satisfies the covariance condition $U\Omega(\rho)U^* = \Omega(U\rho U^*)$ for all unitary $U$. Lemma 2 in the appendix of [20] can therefore be used to show that the minimal CB entropy is achieved when $\gamma_1 = \operatorname{Tr}_2(\mathcal{I}\otimes\Omega)(|\psi\rangle\langle\psi|)$ is the maximally mixed state $\frac{1}{d}\mathrm{I}$ so that $|\psi\rangle$ is maximally entangled and

$$\|\Omega\|_{\mathrm{CB},1\to p} = S(X_\Omega) - \log d \tag{5.3}$$

Moreover, the state $(\mathcal{I}\otimes\Omega_\mu)(|\psi\rangle\langle\psi|)$ has one non-degenerate eigenvalue $\frac{1+(d^2-1)\mu}{d^2}$ and the eigenvalue $\frac{1-\mu}{d^2}$ with multiplicity $d^2-1$. From this one finds

$$\omega_p(\Omega_\mu) = d^{-\frac{p+1}{p}}\Big[(1-\mu+d^2\mu)^p + (d^2-1)(1-\mu)^p\Big]^{1/p} \tag{5.4}$$

and

$$\begin{aligned} S_{\mathrm{CB,min}}(\Omega_\mu) &= -\tfrac{1-\mu}{d^2}\log\tfrac{1-\mu}{d^2} - (d^2-1)\tfrac{1-\mu}{d^2}\log\tfrac{1-\mu}{d^2} - \log d \\ &= \log d - \tfrac{1}{d^2}\big[(1-\mu+d^2\mu)\log(1-\mu+d^2\mu) + (d^2-1)(1-\mu)\log(1-\mu) \end{aligned} \tag{5.5}$$

In the case of qubits, $d = 2$ and (5.4) becomes

$$\|\Omega_\mu\|_{\mathrm{CB},1\to p} = \omega_p(\Omega_\mu) = 2^{-(p+1)/p}\big[(1+3\mu)^p + 3(1-\mu)^p\big]^{1/p} \tag{5.6}$$

which can be compared to

$$\|\Omega_\mu\|_{1\to p} = \nu_p(\Omega_\mu) = 2^{-1}\big[(1+\mu)^p + (1-\mu)^p\big]^{1/p}. \tag{5.7}$$

The strict convexity of $f(x) = x^p$ implies that for $\mu > 0$,

$$(1+\mu)^p = \left(\tfrac{(1+3\mu)+(1-\mu)}{2}\right)^p < \frac{1}{2}\big[(1+3\mu)^p + 3(1-\mu)^p\big]$$

from which it follows that $\|\Omega_\mu\|_{\mathrm{CB},1\to p} > \|\Omega_\mu\|_{1\to p}$. This confirms that, in general, the CB norm $\|\Phi\|_{\mathrm{CB},1\to p}$ of a map $\Phi$ is strictly greater than $\|\Phi\|_{1\to p}$. (This can be seen directly for the identity map $\mathcal{I}$ which corresponds to $\mu = 1$.) For qubits, one can verify explicitly that $S_{\mathrm{CB,min}}(\Phi)$ is achieved with a maximally entangled state and that it decreases monotonically with $\mu$. Numerical work [11] shows that $S_{\mathrm{CB,min}}(\Phi)$ changes from positive to negative at $\mu = 0.74592$, which is also the cut-off for $C_Q(\Phi) = 0$.

The Werner-Holevo channel [52] is $\Phi_{\rm WH}(\rho) = \frac{1}{d-1}\big[(\mathrm{Tr}\,\rho)\mathrm{I} - \rho^T\big]$. One finds that $\gamma_{12}$ has exactly $\binom{d}{2}$ non-zero eigenvalues $\frac{1}{d-1}(a_j^2 + a_k^2)$ with $j < k$ and $a_j^2$ the eigenvalues of $\gamma_1$. One can then use the concavity of $-x\log x$ to show that $S(\gamma_{12}) \geq S(\gamma_1) + \log\frac{d-1}{2}$, which implies that $S_{\rm CB,min}(\Phi_{\rm WH}) = \log\frac{d-1}{2}$ is achieved with a maximally entangled input. Moreover, $S_{\rm CB,min}(\Phi_{\rm WH}) = -1$ for $d = 2$, and $S_{\rm CB,min}(\Phi_{\rm WH}) = 0$ for $d = 3$. One can also use the covariance property $\Phi_{\rm WH}(U\rho U^*) = \overline{U}\Phi_{\rm WH}(\rho)U^T$ and Lemma 2 of [20] to see that $\omega_p(\Phi_{\rm WH})$ is achieved with a maximally entangled state, and verify that

$$\omega_p(\Phi_{\rm WH}) = \big(\tfrac{2}{d-1}\big)^{1-\frac{1}{p}} > \big(\tfrac{1}{d-1}\big)^{1-\frac{1}{p}} = \nu_p(\Phi_{\rm WH}). \tag{5.8}$$

This gives another example for which the CB norm is strictly greater than $\|\Phi\|_{1\to p}$.

However, the CB norm is not always attained on a maximally entangled state. Consider for example the non-unital qubit map $\Phi(\rho) = \lambda\rho + \big(\frac{(1-\lambda)}{2}I + \frac{t}{2}\sigma_3\big)\,\mathrm{Tr}\,\rho$, and the one-parameter family of pure bipartite states $|\psi\rangle_a = \sqrt{a}\,|00\rangle + \sqrt{1-a}\,|11\rangle$ where $0 \leq a \leq 1$. In this case

$$\begin{aligned} \gamma_{12} &= (I\otimes\Phi)(|\psi\rangle_a\langle\psi|) \\ &= \frac{1}{2}\begin{pmatrix} a(1+t+\lambda) & 0 & 0 & 2\lambda\sqrt{a(1-a)} \\ 0 & (1-a)(1+t-\lambda) & 0 & 0 \\ 0 & 0 & a(1-t-\lambda) & 0 \\ 2\lambda\sqrt{a(1-a)} & 0 & 0 & (1-a)(1-t+\lambda) \end{pmatrix} \end{aligned}$$

Numerical computations show that for $p > 1$, $\frac{\|\gamma_{12}\|_p}{\|\gamma_1\|_p}$ is maximized at values $a > 1/2$ when $t > 0$, and values $a < 1/2$ when $t < 0$. Since the state $|\psi\rangle_a$ is maximally entangled only when $a = 1/2$, this demonstrates that the CB norm $\omega_p(\Phi)$ is achieved at a non-maximally entangled state for this family of maps.

## 5.2 Entanglement breaking and preservation

The class of channels for which $(\mathcal{I}\otimes\Phi)(\rho)$ is separable for any input is called entanglement breaking (EB). Those which are also trace preserving are denoted EBT. These maps were introduced in [15] by Holevo who wrote them in the form $\Phi(\rho) = \sum_k R_k \mathrm{Tr}\,\rho E_k$ with each $R_k$ a density matrix and $\{E_k\}$ a POVM, i.e., $E_k \geq 0$ and $\sum_k E_k = I$. They were studied in [19] where several equivalent conditions were proved. The next result shows that EBT channels always have positive minimal CB entropy. Therefore, a channel for which $S_{\rm CB,min}(\Phi)$ is negative always preserves some entanglement.

**Lemma 15** *If $\Phi : M_m \mapsto M_n$ is an EBT map, then for all $p \geq 1$ and positive semi-definite $Q \in M_n \otimes M_m$,*

$$\|(\mathcal{I}_n \otimes \Phi)(Q)\|_p \leq \|\mathrm{Tr}_2\, Q\|_p = \|Q_1\|_p \tag{5.9}$$

**Theorem 16** *If $\Phi$ is an EBT map, then $\omega_p(\Phi) \leq 1$ and $S_{\mathrm{CB,min}}(\Phi)$ is positive.*

Theorem 16 follows immediately from Lemma 15 and Theorem 2 of Section 2.2. The converse does not holds, i.e., $S_{\mathrm{CB,min}}(\Phi) \geq 0$ does not imply that $\Phi$ is EBT. For the depolarizing channel, it is known [43] that $\Omega_\alpha$ is EBT if and only if $|\alpha| \leq \frac{1}{3}$; however, as reported above, $S_{\mathrm{CB,min}}(\Omega_\alpha) > 0$ for $0 < \alpha < 0.74592$. For $d > 3$, the WH channel also has positive CB entropy, although it can not break all entanglement because it is known [52] that $\nu_p(\Phi_{\mathrm{WH}})$ is not multiplicative for sufficiently large $p$.

The proof of Lemma 15 is similar to King's argument [24] for showing multiplicativity of the maximal $p$-norm for EBT maps, and is based on the following inequality due to Lieb and Thirring [33]

$$\mathrm{Tr}\,(C^\dagger D C) \leq \mathrm{Tr}\,(CC^\dagger)^p\, D^p \tag{5.10}$$

for $p \geq 1$ and $D > 0$ positive semi-definite.[1]

**Proof of Lemma 15:** By assumption, we can write $\Phi(\rho) = \sum_k R_k \mathrm{Tr}\, \rho E_k$ with each $R_k$ a density matrix and $\{E_k\}$ a POVM. Then

$$\begin{aligned}(\mathcal{I}_n \otimes \Phi)(Q) &= \sum_{k=1}^{\kappa} [\mathrm{Tr}_2\,(\mathrm{I} \otimes X_k)Q] \otimes R_k \\ &= \sum_{k=1}^{\kappa} G_k \otimes R_k \end{aligned} \tag{5.11}$$

where $G_k = \sum_k [\mathrm{Tr}_2\,(\mathrm{I} \otimes X_k)Q]$. Note that

$$\mathrm{Tr}_2\, Q = \sum_{k=1}^{\kappa} [\mathrm{Tr}_2\,(\mathrm{I} \otimes X_k)Q] = \sum_{k=1}^{\kappa} G_k \tag{5.12}$$

[1]The proof in the Appendix of [33] is based on the concavity of $A \mapsto \mathrm{Tr}\,(BA^{1/m}B)^m$ for $m \geq 1$ and $A, B \geq 0$. This was first proved by Epstein [14]; it is also a special case of Lemma 1.14 in [40], which is proved using complex interpolation in the operator space framework. Araki [1] gave another proof of (5.10), and a simple proof based on Hölder's inequality was given by Simon in Theorem I.4.9 of [48].

With $|e_k\rangle$ the canonical basis in $\mathbf{C}_\kappa$ we define the following matrices in $M_\kappa \otimes M_n \otimes M_n$.

$$R = \sum_k |e_k\rangle\langle e_k| \otimes \mathrm{I}_n \otimes R_k = \begin{pmatrix} \mathrm{I}_n \otimes R_1 & 0 & \dots & 0 \\ 0 & \mathrm{I}_n \otimes R_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \mathrm{I}_n \otimes R_\kappa \end{pmatrix} \tag{5.13}$$

and

$$V = \widetilde{V} \otimes \mathrm{I}_n = \sum_k |e_k\rangle\langle e_1| \otimes G_k^{1/2} \otimes \mathrm{I}_n = \begin{pmatrix} \sqrt{G_1} & 0 & \dots & 0 \\ \sqrt{G_2} & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \sqrt{G_\kappa} & 0 & \dots & 0 \end{pmatrix} \otimes \mathrm{I}_n \tag{5.14}$$

where we adopt the convention of using the subscripts $3, 1, 2$ for $M_\kappa, M_n, M_n$ respectively so that the partial traces $\mathrm{Tr}_1$ and $\mathrm{Tr}_2$ retain their original meaning. It follows that

$$|e_1\rangle\langle e_1| \otimes (\mathcal{I}_n \otimes \Phi)(Q) = V^\dagger R V. \tag{5.15}$$

Applying (5.10) one finds

$$\begin{aligned} \|(\mathcal{I}_n \otimes \Phi)(Q)\|_p^p &= \mathrm{Tr}\,(V^\dagger R V)^p = \mathrm{Tr}_{312}\,(V^\dagger R V)^p \\ &\leq \mathrm{Tr}_{312}\,(VV^\dagger)^p\, R^p \qquad (5.16) \\ &= \sum_k \mathrm{Tr}_{12}\,[(VV^\dagger)^p]_{kk}(\mathrm{I}_n \otimes R_k)^p \\ &= \sum_k \mathrm{Tr}_1\,[(\widetilde{V}\widetilde{V}^\dagger)^p]_{kk} \mathrm{Tr}_2\,(R_k)^p \qquad (5.17) \end{aligned}$$

where $[(\widetilde{V}\widetilde{V}^\dagger)^p]_{kk} = \mathrm{Tr}_3\,(\widetilde{V}\widetilde{V}^\dagger)^p(|e_k\rangle\langle e_k| \otimes \mathrm{I}_n)$ is the $k$-th block on the diagonal of $(\widetilde{V}\widetilde{V}^\dagger)^p$ and $[(VV^\dagger)^p]_{kk} = [(\widetilde{V}\widetilde{V}^\dagger)^p]_{kk} \otimes \mathrm{I}_n$. Since $R_k$ is a density matrix, $\mathrm{Tr}_2\,(R_k)^p \leq 1$. (In fact, we could assume wlog that $R_k = |\theta_k\rangle\langle\theta_k|$ so that $R_k^p = R_k$ and $\mathrm{Tr}_2\,(R_k)^p = 1$.) Therefore,

$$\begin{aligned} \|(\mathcal{I}_n \otimes \Phi)(Q)\|_p^p &\leq \sum_k \mathrm{Tr}_1\,[(\widetilde{V}\widetilde{V}^\dagger)^p]_{kk} \\ &= \mathrm{Tr}_{31}\,(\widetilde{V}\widetilde{V}^\dagger)^p = \mathrm{Tr}_{31}(\widetilde{V}^\dagger\widetilde{V})^p \qquad (5.18) \\ &= \sum_k \mathrm{Tr}_1\, G_k = \mathrm{Tr}_2\, Q \qquad \textbf{QED} \end{aligned}$$

## 5.3 Operational interpretation

Recently Horodecki, Oppenheim and Winter [18] (HOW) obtained results which give an important operational meaning to quantum conditional information, consistent with both positive or negative values. Applying their results to the expression $S_{\rm CB,min}(\Phi) = S(\gamma_{AB}) - S(\gamma_A)$ with $\gamma_{AB} = (\mathcal{I} \otimes \Phi)\big(|\psi\rangle\langle\psi|\big)$ where $|\psi\rangle$ is the minimizer in (1.1) gives the following interpretation:

- A channel for which $S_{\rm CB,min}(\Phi) > 0$ always breaks enough entanglement so that some EPR pairs must be added to enable Alice to transfer her information to Bob.
- A channel for which $S_{\rm CB,min}(\Phi) < 0$ leaves enough entanglement in the optimal state so that some EPR pairs remain after Alice has transferred her information to Bob.

For example, as discussed in Section 5.1 the depolarizing channel is entanglement breaking for $\mu \in [-\frac{1}{3}, \frac{1}{3}]$; for $\mu \in (\frac{1}{3}, 0.74592)$ it always breaks enough entanglement to require input of EPR pairs to transfer Bob's corrupted state back to Alice; and for $\mu > 0.74592$ maximally entangled states retain enough entanglement to allow the distillation of EPR pairs after Bob's corrupted information is transferred to Alice.

Note, however, that the HOW interpretation [18] is an asymptotic result in the sense that it is are based on the assumption of the availability of the tensor product state $\gamma_{AB}^{\otimes n}$ with $n$ arbitrarily large, and is related to the "entanglement of assistance" [49] which is known not to be additive. One would also like to have an interpretation of the additivity of $S_{\rm CB,min}(\Phi)$ so that the "one-shot" formula $-S_{\rm CB,min}(\Phi)$ represents the capacity of an asymptotic process which is not enhanced by entangled inputs. Thus far, the only scenarios for which we have found this to be true seem extremely contrived and artificial.

# 6 Entropy Inequalities

In this section, we show that operator space methods can be used to give a new proof of SSA (1.8). Although the strategy is straightforward, it requires some rather lengthy and tedious bounds on derivatives and norms. Our purpose is not to give another proof of SSA, but to demonstrate the fundamental role of Minkowski-type inequalities and provide some information on the behavior of the $\| \ \|_{(1,p)}$ near $p = 1$.

Differentiation of inequalities of the type found in Section 3.4 often yields entropy inequalities. The procedure is as follows. Consider an inequality of the form $g_L(p) \leq g_R(p)$

valid for $p \geq 1$ which becomes an equality at $p = 1$. Then the function $g(p) = g_R(p) - g_L(p) \geq 0$ for $p \geq 1$ and $g(1) = 0$. This implies that the right derivative $g'(1+) \geq 0$ or, equivalently, that $g'_L(1+) \leq g'_R(1+)$.

Applying this to (3.16) yields

$$-S(Q_1) \leq -S(Q_{12}) + S(Q_2) \tag{6.1}$$

which is the well-known subadditivity inequality $S(Q_{12}) \leq S(Q_1) + S(Q_2)$. Applying the same principle to conjecture (3.17) yields

$$-S(Q_{23}) + S(Q_3) \leq -S(Q_{123}) + S(Q_{13}) \tag{6.2}$$

which is equivalent to strong subadditivity (1.8). (Carlen and Lieb [8] observed that the reverse of (3.17) holds when $t \leq 1$ and used the corresponding left derivative inequality $g'_L(1-) \geq g'_R(1-)$ to obtain another proof of SSA.)

These entropy inequalities can also be obtained by differentiating the corresponding CB Minkowski inequalities (3.13) and (3.14). We will need the following.

**Theorem 17** *For any $X = X_{12}$ in $M_m \otimes M_n$, with $X \geq 0$ and $\operatorname{Tr} X = 1$.*

$$\tfrac{d}{dp} \|X_{12}\|^p_{(1,p)} \Big|_{p=1} = -S(X_{12}) + S(X_1). \tag{6.3}$$

Before proving this result, observe that (3.20) implies $\|W_{12}\|_{(1,p)} = \|W_2\|_p$ and (3.24) implies $\|W_{132}\|_{(1,p,1)} = \|W_{13}\|_{(1,p)}$. Then, when $q = 1$, the inequalities (3.13) and (3.14) imply

$$\|W_2\|^p_p \leq \|W_{12}\|^p_{(1,p)} \tag{6.4}$$

$$\|W_{13}\|^p_{(1,p)} \leq \|W_{123}\|^p_{(1,1,p)}. \tag{6.5}$$

Now, under the assumption that $W_{123} > 0$ and $\operatorname{Tr} W_{123} = 1$, Theorem 17 implies

$$\begin{aligned}
\tfrac{d}{dp} \|W_{21}\|^p_{(1,p)} \Big|_{p=1} &= -S(W_2) \\
\tfrac{d}{dp} \|W_{12}\|^p_{(1,p)} \Big|_{p=1} &= -S(W_{12}) + S(W_1) \\
\tfrac{d}{dp} \|W_{123}\|^p_{(1,1,p)} \Big|_{p=1} &= -S(W_{123}) + S(W_{12}) \\
\tfrac{d}{dp} \|W_{13}\|^p_{(1,p)} \Big|_{p=1} &= -S(W_{13}) + S(W_1)
\end{aligned}$$

Then usual subadditivity and SSA inequalities, (6.1) and (6.2) then follow from the principle, $g'_L(1+) \leq g'_R(1+)$, above and (6.4) and (6.5) respectively.

**Proof of Theorem 17:** The basic strategy is similar to that in Section 2.2, but requires some additional details. Let $X_1 = \mathrm{Tr}_2 X$ and let $Q$ denote the orthogonal projection onto $\ker(X_1)$. Since $QX_1Q = 0$, it follows that $\mathrm{Tr}\,(Q \otimes \mathrm{I}_n)X(Q \otimes \mathrm{I}_n) = 0$. Since $X$ is positive semi-definite this implies that $X = ((\mathrm{I}_m - Q) \otimes \mathrm{I}_n)X((\mathrm{I}_m - Q) \otimes \mathrm{I}_n)$. For fixed $X$ the functions

$$\begin{aligned} v(p,B) &= X^{1/2}(B^{\frac{1}{p}-1} \otimes \mathrm{I}_n)\, X^{1/2}, \quad \text{and} & (6.6)\\ w(p,B) &= X_1^{\frac{1}{2}}\, B^{\frac{1}{p}-1}\, X_1^{\frac{1}{2}}. & (6.7) \end{aligned}$$

are well-defined for $p > 1$, and $B \in \beta(X_1)$ where $\beta(X_1) = \{B \in \mathcal{D} : \ker(B) \subset \ker(X_1)\}$. Since $\|(B^{-\frac{1}{2}(1-\frac{1}{p})} \otimes \mathrm{I}_n)\, X\, (B^{-\frac{1}{2}(1-\frac{1}{p})} \otimes \mathrm{I}_n)\|_p = \|X^{\frac{1}{2}}(B^{-\frac{1}{2}} B^{\frac{1}{p}} B^{-\frac{1}{2}} \otimes \mathrm{I}_n) X^{\frac{1}{2}}\|_p$, it follows from (3.22) and the remarks above that

$$\|X_{12}\|_{(1,p)} = \inf_{B \in \mathcal{D}} \|v(p,B)\|_p = \inf_{B \in \beta(X_1)} \|v(p,B)\|_p. \tag{6.8}$$

The set of density matrices $\mathcal{D}$ is compact, and $\|v(p,B)\|_p$ is bounded below and continuous, hence for each $p > 1$ there is a (i.e., at least one) density matrix $B(p)$ which minimizes $\|v(p,B)\|_p$, so that

$$\|X_{12}\|_{(1,p)} = \|v(p,B(p))\|_p \tag{6.9}$$

Since $p > 1$ and $B(p)$ is a density matrix, $B(p)^{-1+\frac{1}{p}} > \mathrm{I}_m$ which implies

$$v(p,B) \geq X \qquad \text{and} \qquad w(p,B) \geq X_1. \tag{6.10}$$

Furthermore,

$$1 = \mathrm{Tr}\, X \leq \mathrm{Tr}\, v(p,B(p)) \leq (mn)^{\frac{1}{p}-1} \|v(p,B(p))\|_p \tag{6.11}$$

(where the last inequality uses $\|A\|_1 \leq d^{1-\frac{1}{p}} \|A\|_p$ for any positive semi-definite $d \times d$ matrix $A$ and any $p \geq 1$). Replacing $B(p)$ by another density matrix cannot decrease $\|v(p,B(p))\|_p$, hence

$$\|v(p,B(p))\|_p \leq \|v\big(p, \frac{1}{m}\mathrm{I}_m\big)\|_p = m^{1-\frac{1}{p}}\, \|X\|_p \leq m^{1-\frac{1}{p}} \tag{6.12}$$

Combining (6.11) and (6.12) shows that

$$\lim_{p \to 1+} \mathrm{Tr}\,(v(p,B(p)) - X) = 0, \tag{6.13}$$

and, together with (6.10) implies that $v(p, B(p)) \to X$. Also, for any $B \in \beta(X_1)$,

$$\begin{aligned} \operatorname{Tr} v(p,B) &= \operatorname{Tr}_{12} X_{12}\, [B]^{\frac{1}{p}-1} \otimes \mathrm{I}_n) \\ &= \operatorname{Tr} X_1\, [B]^{\frac{1}{p}-1} = \operatorname{Tr} w(p,B) \end{aligned} \tag{6.14}$$

so that $\lim_{p\to 1+} \operatorname{Tr}\left(w(p,B(p)) - X_1\right) = 0$ and $w(p,B(p)) \to X_1$.

Writing out the derivative on the left side of (6.3), we see that we need to show that

$$\lim_{p\to 1+} \frac{1}{p-1}\Big(\operatorname{Tr} v(p,B(p))^p - 1\Big) = -S(X) + S(X_1) \tag{6.15}$$

First note that for $p > 1$.

$$\frac{1}{p-1}\Big(\operatorname{Tr} v(p,B(p))^p - 1\Big) \le \frac{1}{p-1}\Big(\operatorname{Tr} v(p,X_1)^p - 1\Big), \tag{6.16}$$

and a direct calculation shows that the right side of (6.16) converges to $-S(X) + S(X_1)$ as $p \to 1+$. Hence to prove (6.15) it is sufficient to show that

$$\liminf_{p\to 1+} \frac{1}{p-1}\Big(\operatorname{Tr} v(p,B(p))^p - 1\Big) \ge -S(X) + S(X_1) \tag{6.17}$$

Hölder's inequality implies

$$\begin{aligned} 1 = \|X_1\|_1 &= \|B(p)^{\frac{1}{2}-\frac{1}{2p}} \Big(B(p)^{\frac{1}{2p}-\frac{1}{2}}\, X_1\, B(p)^{\frac{1}{2p}-\frac{1}{2}}\Big) B(p)^{\frac{1}{2}-\frac{1}{2p}}\|_1 \\ &\le \|B(p)^{\frac{1}{2}-\frac{1}{2p}}\|^2_{2p/p-1}\ \|B(p)^{\frac{1}{2p}-\frac{1}{2}}\, X_1\, B(p)^{\frac{1}{2p}-\frac{1}{2}}\|_p \\ &= \|w(p,B(p))\|_p \end{aligned} \tag{6.18}$$

Combining this with (6.14) gives a bound on the numerator on the left in (6.15)

$$\begin{aligned} \operatorname{Tr} v(p,B(p))^p - 1 &\ge \operatorname{Tr} v(p,B(p))^p - \operatorname{Tr} v(p,B(p)) \\ &\quad -\Big[\operatorname{Tr} w(p,B(p))^p - \operatorname{Tr} w(p,B(p))\Big] \end{aligned} \tag{6.19}$$

The mean value theorem for the function $g(p) = x^p$ implies that for some $p_1, p_2 \in [1,p]$

$$\frac{1}{p-1}\Big(\operatorname{Tr} v(p,B(p))^p - \operatorname{Tr} v(p,B(p))\Big) = \operatorname{Tr} v(p,B(p))^{p_1}\, \log v(p,B(p)) \tag{6.20a}$$

$$\frac{1}{p-1}\Big(\operatorname{Tr} w(p,B(p))^p - \operatorname{Tr} w(p,B(p))\Big) = \operatorname{Tr} w(p,B(p))^{p_2}\, \log w(p,B(p)) \tag{6.20b}$$

The convergence in (6.13) and following (6.14) imply

$$\begin{aligned} \lim_{p\to 1} \operatorname{Tr} v(p, B(p))^{p_1} \log v(p, B(p)) &= -S(X) && (6.21a) \\ \lim_{p\to 1} \operatorname{Tr} w(p, B(p))^{p_2} \log w(p, B(p)) &= -S(X_1). && (6.21b) \end{aligned}$$

Combining (6.20), (6.21)and (6.19) gives (6.17). **QED**

**Remark:** The proof above relies on the convergence of $\lim_{p\to 1+} X^{\frac{1}{2}}(B^{\frac{1}{p}-1} \otimes \mathrm{I}_n)\, X^{\frac{1}{2}} = X$ and $\lim_{p\to 1+} X^{\frac{1}{2}} X_1^{\frac{1}{2}} B^{\frac{1}{p}-1}\, X_1^{\frac{1}{2}} = X_1$, but tells us nothing at all about the behavior of $B(p)$ as $p \to 1+$. By making a few changes at the end of this proof and exploiting Klein's inequality, we can also show that $\lim_{p\to 1+} B(p) = X_1$.

Klein's inequality [28, 37] states that

$$\operatorname{Tr} A \log A - \operatorname{Tr} A \log B \geq \operatorname{Tr}(A - B) \tag{6.22}$$

with equality in the case $\operatorname{Tr} A = \operatorname{Tr} B$ if and only if $A = B$.

Now, replace (6.19) by

$$\operatorname{Tr} v(p, B(p))^p - 1 \;=\; \operatorname{Tr} v(p, B(p))^p - \operatorname{Tr} v(p, B(p)) + (\operatorname{Tr} w(p, B(p)) - 1). \tag{6.23}$$

Then use the mean value theorem for the function $g_2(p) = y^{\frac{1}{p}}$ to replace (6.20b) by

$$\frac{1}{p-1}\Big(\operatorname{Tr} w(p, B(p)) - 1\Big) \;=\; -\frac{1}{\widetilde{p}^2} \operatorname{Tr} X_1^{\frac{1}{2}}\, B(p)^{\frac{1}{\widetilde{p}}-1)} \log B(p)\, X_1^{\frac{1}{2}}. \tag{6.24}$$

We could use (6.22) with $A = B(p)^{-\frac{1}{2}(1-\frac{1}{\widetilde{p}})} X_1 B(p)^{-\frac{1}{2}(1-\frac{1}{\widetilde{p}})}$ together with the fact that $A$ and $w(\widetilde{p}, B(p))$ have the same non-zero eigenvalues to bound the right side of (6.24) below by $-\frac{1}{\widetilde{p}^2} S[w(\widetilde{p}, B(p)] + \operatorname{Tr} w(\widetilde{p}, B(p)) - 1$. However, because $1 < \widetilde{p} < p$ implies $B^{1/\widetilde{p}} > B^{1/p}$, we cannot extend (6.12) and (6.14) to conclude that this converges to $S(X_1)$.

Instead, we first observe that the compactness of the set of density matrices $\mathcal{D}$ implies that we can find a sequence $p_k \to 1+$ such that $\|X_{12}\|_{(1,p)} = \|v(p_k, B(p_k))\|_{p_k}$ and $B_k \to B^* \in \mathcal{D}$. If $B^*$ is not in $\beta(X_1)$, then the right side of the first line of (6.24) $\to +\infty$ giving a contradiction with (6.16). Hence $B^* \in \beta(X_1)$. Therefore, (6.24) and (6.22) imply

$$\lim_{p_k\to\infty} \frac{1}{p_k - 1}\Big(\operatorname{Tr} w(p_k, B(p_k)) - 1\Big) \;=\; -\operatorname{Tr} X_1 \log B^* \geq S(X_1). \tag{6.25}$$

Inserting this in (6.23) yields

$$\begin{aligned} \lim_{p_k\to\infty} \frac{1}{p_k - 1}\Big(\operatorname{Tr} v(p_k, B(p_k))^p - 1\Big) &= -S(X_{12}) - \operatorname{Tr} X_1 \log B^* \\ &\geq -S(X_{12}) + S(X_1). && (6.26) \end{aligned}$$

Combining these results with (6.16), we conclude that equality holds in (6.26) and that

$$- \operatorname{Tr} X_1 \log B^* = S(X_1) = -\operatorname{Tr} X_1 \log X_1. \tag{6.27}$$

We can now use the condition for equality in (6.22) to conclude that $B^* = X_1$. Since this is true for the limit of *any* convergent sequence of minimizers $B(p_k)$ with $p_k \to 1$, we have also proved the following which is of independent interest.

**Corollary 18** *For $X \in M_m \otimes M_n$ with $X \geq 0$ and $\operatorname{Tr} X = 1$ and $p \in (1, 2]$, let $B(p) \in \mathcal{D}$ minimize $\|X\|_{(1,p)}$, i.e., $\|X^{\frac{1}{2}} (B^{\frac{1}{p}-1} \otimes \mathrm{I}_n)\, X^{\frac{1}{2}}\|_p = \|X\|_{(1,p)}$. Then $\lim\limits_{p \to 1+} B(p) = X_1 \equiv \operatorname{Tr}_2 X$.*

**Acknowledgments:** The work of M.J. was supported in part by National Science Foundation grant DMS-0301116. The work of C.K. was supported in part by the National Science Foundation under grant DMS-0400426. The work of M.B.R. was supported in part by the National Security Agency (NSA) and Advanced Research and Development Activity (ARDA) under Army Research Office (ARO) contract number DAAD19-02-1-0065; and by the National Science Foundation under Grant DMS-0314228.

This work had its genesis in a workshop in 2002 at the Pacific Institute for the Mathematical Sciences at which M.J. and M.B.R. participated. Part of this work was done while I.D. and M.B.R. were visiting the Isaac Newton Institute. The authors are grateful to these institutions for their hospitality and support. Finally, we thank Professor Andreas Winter for discussions about possible interpretations of $S_{\mathrm{CB,min}}(\Phi)$.

# A Purification

To make this paper self-contained and accessible to people in fields other than quantum information we summarize the results needed to prove Lemma 4.

Any density matrix in $\mathcal{D}_d$ can be written in terms of its spectral decomposition (restricted to $[\ker(\gamma)]^\perp$) as $\gamma = \sum_{k=1}^m \lambda_k \, |\phi_k\rangle\langle\phi_k|$ where each eigenvalue $\lambda_k > 0$ and counted in terms of its multiplicity so that the eigenvectors $\{|\phi_k\rangle\}$ are orthonormal. If we then let $\{|\chi_k\rangle\}$ be any orthonormal basis of $\mathbf{C}^m$ and define $|\Psi\rangle \in \mathbf{C}^d \otimes \mathbf{C}^m$ as

$$|\Psi\rangle = \sum_{k=1}^m \sqrt{\lambda_k} \, |\phi_k \otimes |\chi_k\rangle. \tag{A.1}$$

then $\gamma = \operatorname{Tr}_2 |\Psi\rangle\langle\Psi|$ and (A.1) is called a *purification* of $\gamma$.

Conversely, given a normalized vector $|\Psi\rangle \in \mathbf{C}^n \otimes \mathbf{C}^m$, it is a straightforward consequence of the singular value decomposition that $|\Psi\rangle$ can be written in the form

$$|\Psi\rangle = \sum_k \mu_k \, |\phi_k \otimes |\chi_k\rangle \tag{A.2}$$

with $\{|\phi_k\}$ and $\{|\chi_k\rangle\}$ orthonormal sets in $\mathbf{C}^n$ and $\mathbf{C}^m$ respectively. (This is often called the "Schmidt decomposition" in quantum information theory. For details and some history see Appendix A of [25].) It follows from (A.2) that the reduced density matrices $\gamma_1 = \mathrm{Tr}_2|\Psi\rangle\langle\Psi|$ and $\gamma_2 = \mathrm{Tr}_1|\Psi\rangle\langle\Psi|$ have the same non-zero eigenvalues. Although our interest here is for $\mathcal{H} = \mathbf{C}^m$, these results extend to infinite dimensions and yield the following

**Corollary 19** *When $|\Psi_{AB}\rangle$ is a bipartite pure state in $\mathcal{H}_A \otimes \mathcal{H}_B$, then its reduced density matrices $\gamma_A = \mathrm{Tr}_B|\Psi\rangle\langle\Psi|$ and $\gamma_B = \mathrm{Tr}_A|\Psi\rangle\langle\Psi|$ have the same entropy, i.e., $S(\gamma_A) = S(\gamma_B)$.*

# References


[1] H. Araki, "On an inequality of Lieb and Thirring *Lett. Math. Phys.* **19**, 167–170 (1990).

[2] K.M.R. Audenaert, "A note on the $p \to q$ norms of completely positive maps" math-ph/0505085

[3] G. G. Amosov, A. S. Holevo, and R. F. Werner, "On Some Additivity Problems in Quantum Information Theory", *Problems in Information Transmission*, **36**, 305–313 (2000). math-ph/0003002.

[4] H. Barnum, M.A. Nielsen, and B. Schumacher, "Information transmission through a noisy quantum channel" *Phys. Rev. A*, **57**, 4153–4175 (1998).

[5] C. H. Bennett, P.W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted classical capacity of noisy quantum channels" *Phys.Rev.Lett.* **83**, 3081–84 (1999) quant-ph/9904023

[6] C. H. Bennett, P.W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem" *IEEE Trans. Inform. Theory* **48**, 2637–2655 (2002). quant-ph/0106052

[7] D.P. Blecher and V. I. Paulsen, "Tensor products of operator spaces" *J. Funct. Anal.* **99** 262–292 (1991).

[8] E. Carlen and E. Lieb, "A Minkowski type trace inequality and strong subadditivity of quantum entropy" *Amer. Math. Soc. Transl.* **189**, 59–62 (1999). Reprinted in [31].

[9] M-D Choi, "Completely Positive Linear Maps on Complex Matrices" *Lin. Alg. Appl.* **10**, 285–290 (1975).

[10] I. Devetak, "The Private Classical Capacity and Quantum Capacity of a Quantum Channel" *IEEE Trans. Inform. Theory* **51**, 44–55 (2005). quant-ph/0304127

[11] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin "Quantum-channel capacity of very noisy channels" *Phys. Rev. A* **57**, 830–839 (1998); erratum **59**, 1717 (1999). quant-ph/9706061

[12] E. G. Effros and Z. J. Ruan, "Self-duality for the Haagerup tensor product and Hilbert space factorizations" *J. Funct. Anal.* **100** 257–284 (1991).

[13] E. G. Effros and Z. J. Ruan, *Operator Spaces* (Oxford Univ. Press, 2000).

[14] H. Epstein, "Remarks on two theorems of E. Lieb" *Commun. Math. Phys.* **31**, 317–325 (1973).

[15] A. S. Holevo, "Coding Theorem for Quantum Channels" quant-ph/9809023; "Quantum coding theorems", *Russian Math. Surveys*, **53**, 1295–1331 (1999).

[16] A. S. Holevo "On Entanglement-Assisted Classical Capacity" *J. Math. Phys.* **43**, 4326–4333 (2002). quant-ph/0106075

[17] A. S. Holevo and R.F. Werner, "Evaluating capacities of bosonic Gaussian channels" *Phys. Rev. A* **63**, 032312 (2001).

[18] M.Horodecki, J. Oppenheim and Andreas Winter "Quantum information can be negative" quant-ph/0505062

[19] M.Horodecki, P. Shor, and M. B. Ruskai "Entanglement Breaking Channels" *Rev. Math. Phys* **15**, 629–641 (2003). (quant-ph/030203)

[20] A. Jencŏvá , "A relation between completely bounded norms and conjugate channels" quant-ph/0601071

[21] M. Junge, "Factorization theory for spaces of operators" Habilitation thesis Kiel University (1996).

[22] M. Junge, "Vector-valued $L_p$ spaces for von Neumann algebras with QWEP"

[23] M. Junge and Z.-J. Ruan: "Decomposable Maps on Non-commutative $L_p$ spaces," *Contemporary Mathematics* **365**, 355–381.(2004).

[24] C. King, "Maximal p-norms of entanglement breaking channels", *Quantum Information and Computation*, **3**, no. 2, 186–190 (2003).

[25] C. King and M.B. Ruskai "Minimal Entropy of States Emerging from Noisy Quantum Channels" *IEEE Trans. Info. Theory* **47**, 1–19 (2001).

[26] C.King and M. B. Ruskai, "Comments on multiplicativity of maximal p-norms when p = 2" in *Quantum Information, Statistics and Probability* ed. by O. Hirota, 102–114 (World Scientific, 2004) quant-ph/0401026.

[27] A. Kitaev, *Classical and Quantum Computation* (AMS, 2002).

[28] O. Klein, "Zur quantenmechanischen begründung des des zweiten haupsatzes der wärmelehre " *Zeit. für Physik* **72**, 767-775 (1931).

[29] K. Kraus, "General state changes in quantum theory" *Ann. Phys.* **64**, 311–335 (1971).

[30] K. Kraus, *States, Effects and Operations: Fundamental Notions of Quantum Theory* (Springer-Verlag, 1983).

[31] *Inequalities: Selecta of E. Lieb*, M. Loss and M.B. Ruskai, eds (Springer, 2002).

[32] E.H. Lieb and M.B. Ruskai, "Proof of the Strong Subadditivity of Quantum Mechanical Entropy" *J. Math. Phys.* **14**, 1938–1941 (1973). Reprinted in [31].

[33] E. Lieb and W. Thirring, "Inequalities for the Moments of the Eigenvalues of the Schrödinger Hamiltonian and Their Relation to Sobolev Inequalities", in *Studies in Mathematical Physics*, E. Lieb, B. Simon, A. Wightman eds., pp. 269–303 (Princeton University Press, 1976). Reprinted in [31].

[34] S. Lloyd, "Capacity of a noisy quantum channel" *Phys. Rev. A* **55**, 1613–1622 (1997).

[35] E. Nelson, "Notes on non-commutative integration" *J. Func. Anal.* **15** 103–116 (1974).

[36] M. Nielsen, *Quantum Information Theory*, PhD thesis, Univ. of New Mexico, 1998.

[37] M. Nielsen and I. Chuang *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[38] V. Paulsen, *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, 2000).

[39] G. Pisier, *The Operator Hilbert Space OH, Complex Interpolation and Tensor Norms Memoirs AMS*, **122** (American Mathematical Society, 1996).

[40] G. Pisier, "Non-Commutative Vector Valued $L_p$-spaces and Completely $p$-summing Maps" (Société Mathématique de France, 1998).

[41] G. Pisier, *Introduction to Operator Space Theory* (Cambridge University Press, 2003).

[42] M.B. Ruskai, "Inequalities for Quantum Entropy: A Review with Conditions for Equality" *J. Math. Phys.* **43**, 4358–4375 (2002); erratum **46**, 019901 (2005).

[43] M. B. Ruskai, "Qubit Entanglement Breaking Channels" *Rev. Math. Phys.* **15**, 643-662 (2003). (quant-ph/0302032)

[44] Z-J. Ruan "Subspaces of $C^*$-algebras" *J. Funct. Anal.* **76**, 217–230 (1988).

[45] I. E. Segal, "A non-commutative extension of abstract integration" *Ann. of Math.* **57**, 401–457 (1953).

[46] P. W. Shor, announced at MSRI workshop, (November, 2002). Notes are available at www.msri.org/publications/ln/msri/2002/quantumcrypto/shor/1/index.html;.

[47] P. W. Shor, "Equivalence of Additivity Questions in Quantum Information Theory", *Commun. Math. Phys.* **246**, 453–472 (2004). quant-ph/0305035

[48] B. Simon *The Statistical Mechanics of Lattice Gases* (Princeton Univ. Press, 1993).

[49] J. Smolin, F. Verstraete and A. Winter "Entanglement of assistance and multipartite state distillation" quant-ph/0505038

[50] W.F. Stinespring, "Positive functions on $C^*$-algebras" *Proc. Amer. Math. Soc.* **6**, 211–216 (1955).

[51] J. Watrous, "Notes on super-operator norms induced by Schatten norms" *Quantum Inf. Comput.* **5** 57-67 (2005). quant-ph/0411077

[52] R. F. Werner and A. S. Holevo, "Counterexample to an additivity conjecture for output purity of quantum channels", *Jour. Math. Phys.* **43**, no. 9, 4353–4357 (2002).